To the Graduate Council:

I am submitting herewith a dissertation written by Hem Raj Joshi entitled "Optimal Control Problems in PDE and ODE Systems". I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mathematics.

Suzanne Lenhart, Major Professor

We have read this thesis
and recommend its acceptance:

Xiaobing Feng

G. Samuel Jordan

John D. Landes

Accepted for the Council:

Dr. Anne Mayhew

Vice Provost and Dean of
Graduate Studies

(Original signatures are on file in the Graduate Studies Office.)

Optimal Control Problems in PDE and ODE Systems

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Hem Raj Joshi
May 2002

Dedication

This thesis is dedicated to my parents

Khem Raj Joshi and Haripriya Joshi

in appreciation for their encouragement

throughout my life

Acknowledgements

I would like to express my most sincere gratitude to my advisor, Professor Suzanne Lenhart. Without her help I would not be able to complete my study. Her patience, insight, understanding, and friendship through my graduate career have been greatly appreciated.

I would also like to thank my Dissertation Committee: Professor Xiaobing Feng, Professor G. Samuel Jordan, and Professor John D. Landes. In particular, I would like to thank Professor Feng and Professor Jordan who have taught me a great deal of applied mathematics in their courses over the past years.

I would like to thank the concerned authorities at the Tribhuban University, Kathmandu, Nepal for providing a long study leave to complete my dissertation.

I would like to thank my wife Dr. Rashika Joshi, son Umang Raj Joshi, daughter Imani Joshi and my family for their help and steadfast support as I pursued my doctoral degree. Finally, I would like to thank all my friends, fellow graduate students and Nepali community in Knoxville who, along with my family and advisors, have given me the encouragement and help I needed to reach my goals.

Abstract

This dissertation contains three separate optimal control problems involving partial differential equations (PDEs) or ordinary differential equations (ODEs). In each problem, an objective functional representing the goal of the control process is minimized. First, a system of ordinary differential equations which describe the interaction of Human Immunodeficiency Virus (HIV) and $CD4^+T$-cells in the human immune system is studied. Two controls representing drug treatment strategies of this model are explored. Existence and uniqueness results for the optimal control pair are established. The optimality system is derived and then solved numerically using an iterative method with the Runge-Kutta fourth order scheme.

Second, an unknown coefficient of the interaction term of a parabolic system with a Neumann boundary condition in a multi-dimensional bounded domain is identified. The solution of this system represents the concentrations of predator and prey populations. Given partial (perhaps noisy) observations of the true solution in a subdomain, we seek to "identify" the coefficient of the interaction term using an optimal control problem technique. This method of solving this identification problem is based on Tikhonov's regularization and the optimal control for a fixed regularization parameter represents the approximate solution of the inverse problem. The existence and uniqueness of the optimal control are established, and an optimality system is derived. As the regularization parameter goes to zero, the identification problem is solved, and an example illustrating how to find a solution numerically is presented.

Third, a problem involving optimal control of a convective velocity coefficient depending on space and time in a parabolic equation is treated. This work applies to a one dimensional fluid flow through a soil-packed tube in which a contaminant is initially distributed. The existence of an optimal control and an optimality system are derived. This problem requires more regularity on the control set which results in a PDE characterization of an optimal control.

Contents

List of Figures

Part I

Introduction

1. Background

Optimal control theory, developed by Lev Semenovic Pontryagin and his co-workers in the late 1950s, generalizes the calculus of variations by extending its range of applicability. Pontryagin developed the maximum principle for the optimal control of finite dimensional problems governed by ordinary differential equations (ODEs). In optimal control problems, variables are divided into two classes, **state** variables and **control** variables. The movement of the state variables is governed by first order differential equations. The trajectories of the state variables are influenced directly by the control variables. It is important to note that the number of control variables may not match with the number of state variables [**1, 5**]. The control variables and thus the state variables are chosen to maximize an objective functional, respecting the desired goal.

As an illustration of a simple version of Pontryagin's Maximum Principle, the following problem is considered [**5**]. Given a piecewise continuous control function $u(t)$, there is an associated continuous and piecewise differentiable state $x(t)$ defined on $[t_0, t_1]$, a finite time interval, that solves

$$x'(t) = g(t, x(t), u(t))$$

with initial conditions

$$\begin{aligned} &x(t_0) = x_0 && \text{fixed} \\ &x(t_1) && \text{free} \end{aligned}$$

Our goal is to find u^* that maximizes the objective functional

$$J(u^*) = \max_u J(u), \tag{1.1}$$

where $J(u) = \displaystyle\int_{t_0}^{t_1} f(t, x(t), u(t))\, dt$. Minimization problems can also be treated. Here f and g are assumed to be continuous differentiable functions of three independent arguments.

The following result gives the necessary conditions for an optimal control [**1, 5**].

THEOREM 1.1. *If $u^*(t)$ is an optimal control in (1.1), with $x^*(t)$ being the corresponding state, then it is necessary that there exists a differentiable adjoint function λ, where for all $t_0 \leq t \leq t_1$, $\lambda(t) \neq 0$, and*

$$H(t, x^*(t), u(t), \lambda(t)) \leq H(t, x^*(t), u^*(t), \lambda(t)),$$

for all controls $u(t)$, where the Hamiltonian is defined by

$$H(t, x(t), u(t), \lambda(t)) = f(t, x(t), u(t)) + \lambda g(t, x(t), u(t)).$$

and $\lambda(t)$ satisfies

$$\lambda' = -\frac{\partial H}{\partial x}$$

and the transversality condition, $\lambda(t_1) = 0$.

2. Bounded Controls in ODE case

If the controls are bounded [**1**],we consider

$$J(u^*) = \max_u \int_{t_0}^{t_1} f(t, x(t), u(t))\, dt$$

subject to $x'(t) = g(t, x(t), u(t)), x(t_0) = x_0$ and $a \leq u(t) \leq b$.

We assume that H is a nonlinear differentiable function of u.

We need to adjust the optimal control as follows:

$$\begin{aligned} f_u + \lambda g_u &\leq 0 \quad \text{at } u^*(t) \quad \Rightarrow \quad u^*(t) = a \\ f_u + \lambda g_u &= 0 \quad \text{at } u^*(t) \quad \Rightarrow \quad a \leq u^*(t) \leq b \\ f_u + \lambda g_u &\geq 0 \quad \text{at } u^*(t) \quad \Rightarrow \quad u^*(t) = b. \end{aligned} \tag{2.2}$$

The other necessary conditions for the optimal control are:

$$\begin{array}{lll} -\frac{\partial H}{\partial x} = \lambda' & \text{i.e. } \lambda'(t) = -(f_x + g_x) & \text{adjoint equation } + \text{ B. C.} \\ \frac{\partial H}{\partial \lambda} = x' & \text{i.e. } x'(t) = g & \text{state equation.} \end{array}$$

We obtain the optimality condition with the help of the Lagrangian, which is formed by appending the constraints to the objective with multiplier functions w_1, w_2. So, $L =$

$f(t, x(t), u(t)) + \lambda g(t, x(t), u(t)) + w_1(b - u) + w_2(u - a)$ gives the necessary conditions for a constrained maximum/minimum problem with respect to u:

$$\begin{aligned} &\tfrac{\partial L}{\partial u} = f_u + \lambda g_u - w_1 + w_2 = 0, \\ &w_1(t) \geq 0 \qquad w_1(b - u^*) = 0, \\ &w_2(t) \geq 0 \qquad w_2(u^* - a) = 0. \end{aligned} \tag{2.3}$$

If $u^*(t) = b$, then $u^*(t) - a > 0$, so $w_2(t) = 0$. Hence $f_u + \lambda g_u - w_1 = 0$ at t. But we have $w_1(t) \geq 0$; which requires $f_u + \lambda g_u \geq 0$ at t. The other possibilities in (2.2) are also verified similarly from these Lagrangian conditions.

3. Optimal Control of PDEs

The theory of optimal control of systems governed by partial differential equations (PDEs) was developed by J. L. Lions [**2**]. Even though other similarities occur, there is no full generalization of Pontryagin's maximum principle to nonlinear PDEs. Indeed, the ideas of Pontryagin's Maximum Principle are used to aid in characterizing an optimal control through an optimality system, which involves the state system coupled with the adjoint system. Explicitly, there are several connections between finite and infinite dimensional optimal control theory. Given the optimal controls and the corresponding state variables, there exist adjoint variables that satisfy systems in which the source terms of the adjoint PDEs equal the partial derivatives of the integrand of the objective functional with respect to the state variables. In the case that the state equation is parabolic, then the adjoint equation is a backward parabolic equation and a transversality condition occurs at the final time. See the book by Li and Yong for some known cases of Pontryagin's Maximum Principle for PDEs [**4**].

The typical setting of an optimal control problem with a single PDE state equation is as follows [**2**]:

- A control h belongs to some set U, the set of admissible controls which is at our disposition.

- The state $u(h)$ of the system to be controlled, depending on h, is the solution of the equation

  $$\mathcal{L}u(h) = F(h, u, x, t)$$

  with appropriate boundary conditions, where $\mathcal{L}$ is known PDE operator, which specifies the system to be controlled.
- The "objective functional" $J(h)$ to be maximized or minimized is defined in terms of $u(h)$ and h:

  $$J(h) = \int_Q f(x, t, u(h), h)\, dx\, dt.$$

  The definitive goal is to find $h^* \in U$ such that

  $$J(h^*) = \inf J(h) \text{ or } \sup J(h)$$

  over all $h \in U$.

The following is an outline of the solution technique. If the state system is nonlinear, the existence and uniqueness of the state system are proved via an iterative scheme. The existence of an optimal control is obtained through a maximizing or minimizing sequence argument. The optimality system consists of the state system coupled with the adjoint system together with the characterization of the optimal control. To obtain the necessary conditions for the optimal control, the objective functional is differentiated with respect to the control. We compute the directional derivative (Gateaux derivative) of the objective functional with respect to h in direction $k \in U$, i.e.,

$$\lim_{\varepsilon \to 0^+} \frac{J(h + \varepsilon k) - J(h)}{\varepsilon}.$$

The limit is denoted as $\dfrac{\partial J}{\partial h}(h; k)$, where k denotes the direction. Since the objective functional usually contains the state variable u, thus it also needs to be differentiated with respect to the control. The difference quotients $\dfrac{u(h + \varepsilon k) - u(h)}{\varepsilon}$ are shown to converge to a sensitivity function, ψ. A *priori* estimates are needed for the existence of the state system, the existence of an optimal control, and the convergence of the difference quotients to ψ. The function ψ solves a linearized version of the state equation. Finally, the adjoint of the operator of the linearized ψ equation is found. The components of the right-hand side of the adjoint system are the derivatives of the integrand of the objective functional with respect to each state. To compute the optimal control h^* from $\dfrac{\partial J}{\partial h}$, the adjoint system is introduced with a transversality condition and other appropriate boundary conditions. The explicit form of the optimal control is determined using standard techniques involving choices of the variation function k. It is important to note that, if the objective functional has partial derivatives of the control in its integrand then the optimal control satisfies a PDE (i. e., an explicit expression is not obtained) . In the parabolic case, uniqueness for the optimality system, which characterizes the unique optimal control, holds only for sufficiently small time. This condition is due to the opposite time orientations of the state and adjoint systems.

In part II, we consider optimal control of an HIV immunology model:

$$\frac{dT(t)}{dt} = s_1 - \frac{s_2 V(t)}{B_1 + V(t)} - \mu T(t) - kV(t)T(t) + u_1(t)T(t) \tag{3.4}$$

$$\frac{dV(t)}{dt} = \frac{g(1-u_2(t))V(t)}{B_2 + V(t)} - cV(t)T(t) \tag{3.5}$$

satisfying $T(0) = T_0$ and $V(0) = V_0$, where T, V represent the concentrations of the uninfected $CD4^{+}T$ cells and the free infectious virus particles respectively, and u_1, u_2 represent controls for two different treatment strategies. The term $s_1 - \frac{s_2 V(t)}{B_1+V(t)}$ is the source/proliferation of uninfected T cells , $-\mu T(t)$ is the natural loss of uninfected T cells, $-kV(t)T(t)$ is loss by infection, $\frac{gV(t)}{B_2+v(t)}$ is viral production in plasma and $-cV(t)T(t)$ is the viral loss. Similarly, μ is death rate of T cells, k is infection rate of T cells, g is the input rate of external virus source, c is loss rate of virus and B_1, B_2 are half saturation constants.

Our goal is to boost the immune system using the control u_1 and delay the HIV progression using another control u_2. The objective functional to be maximized is

$$J(u_1, u_2) = \int_0^{t_f} [T - (A_1 u_1^2 + A_2 u_2^2)]\, dt. \tag{3.6}$$

The first term represents benefit of T cells and the other terms are systemic costs of the drug treatments, where t_f is the final time. The positive constants A_1 and A_2 balance the size of the terms, u_1^2, u_2^2, and reflect the severity of the side effects of the drugs.

We seek an optimal control pair, u_1^*, u_2^* such that

$$J(u_1^*, u_2^*) = \max\{J(u_1, u_2) | (u_1, u_2) \in U\}$$

where $U = \{(u_1, u_2) | u_i \text{ measurable },\ a_i \le u_i \le b_i, \quad t \in [0, t_f],\ \text{ for } i = 1, 2\}$ is the control set, and a_i, b_i are lower and upper bounds on controls.

The existence and uniqueness results for the optimal control pair are established. We further derive the optimality system and then solve it numerically using an iterative method with Runge-Kutta fourth order scheme.

In part III, we consider a parabolic inverse problem of identification type:

$$\begin{aligned} u_t - b_1 \Delta u &= (a_1 - du)u - h(x)uv \quad \text{in } Q = \Omega \times (0, T) \\ v_t - b_2 \Delta v &= -a_2 v + h(x)uv \end{aligned} \tag{3.7}$$

with initial and boundary conditions:

$$\begin{aligned} u(x, 0) &= u_0(x), \quad v(x, 0) = v_0(x) \qquad && \text{for } x \in \Omega \\ \frac{\partial u}{\partial \nu} &= 0, \quad \frac{\partial v}{\partial \nu} = 0 && \text{on } \partial\Omega \times (0, T). \end{aligned}$$

The coefficients and terms can be interpreted as :

$h(x)$	$=$ coefficient of the interaction term to be identified.
u	$=$ prey population (state variable).
v	$=$ predator population (state variable).
b_1, b_2	$=$ diffusion coefficients
a_1, a_2, d	$=$ standard logistic growth terms.

Our class of admissible interaction coefficients is

$$U = \{h \in L^2(\Omega) : 0 \le h \le M\}$$

We seek to minimize the objective functional:

$$J(h) = \int_W [(u(h) - z_1)^2 + (v(h) - z_2)^2] \, dx \, dt$$

over $h \in U$, where $W \subset Q$, the measure of W is positive, and z_1, z_2 are observations of the prey and predator populations. We denote the dependence of u and v on the interaction coefficient by

$$u = u(h), \quad v = v(h),$$

Given partial (and perhaps noisy) observations z_1, z_2 of the true solution in a subdomain W of Q, we seek to "identify" $h(x)$.

To identify h, treat h as a control and choose h such that corresponding u, v are close to given observations z_1, z_2, and an approximate objective functional is minimized. The approximate objective functional for $\beta > 0$ is:

$$J_\beta(h) = \frac{1}{2} \int_W [(u - z_1)^2 + (v - z_2)^2] \, dx \, dt + \frac{\beta}{2} \int_\Omega h(x)^2 \, dx \,. \tag{3.8}$$

The second term in the right hand side (RHS) is "artificial". This method of solving these identification problems is based on Tikhonov's regularization [**6**]. Tikhonov's approach and its variants seek to minimize the functional $J_\beta(h)$ for a fixed β over a set of unknown data. The function h_β which achieves the minimization of $J_\beta(h)$ represents an approximate solution of the inverse problem.

For each $\beta > 0$, one considers the unknown data, h, as a control which belongs to U; the control has to be adjusted in order to minimize the functional $J_\beta(h)$. We seek to find the minimum of the approximate objective functional over h in U :

$$J_\beta(h_\beta) = \inf_{h \in U} J_\beta(h).$$

Letting the sequence of β tend toward zero, we verify that the corresponding sequence, h_β, converges in an appropriate sense to the desired coefficient.

The existence and uniqueness of the optimal control approximating the desired coefficient are obtained, an optimality system is derived, the identification problem is discussed and an example illustrating how to find a solution numerically is presented.

In part IV, we consider optimal control of the coefficient of the first spatial derivative in a parabolic equation. We treat the case of one spatial dimension and the control depends on space and time. Control of such a coefficient requires more regularity on the control set than controlling the coefficient of a zero$^{\text{th}}$ order term, and it lead to a partial differential equation (PDE) characterization of an optimal control instead of the usual explicit characterization.

We consider the parabolic equation (in non-divergent form) is

$$\begin{aligned} u_t(x,t) &= au_{xx} - h(x,t)u_x + f \quad && \text{in } Q = (0,L) \times (0,T) \\ u(x,0) &= u_0(x) && 0 < x < L,\ t = 0, \end{aligned} \tag{3.9}$$

with the boundary conditions

$$\begin{aligned} u(0,t) = 0 \qquad &\text{for all} \quad t \in (0,T) \\ u_x(L,t) = 0 \qquad &\text{for all} \quad t \in (0,T). \end{aligned}$$

The convective velocity coefficient $h(x,t)$ is taken to be the control.

The objective functional is

$$J(h) = \alpha \int_0^T \int_0^L u(x,t)\,dx\,dt + \frac{\beta}{2} \int_0^T \int_0^L \left[h^2(x,t) + h_t^2(x,t) + h_x^2(x,t)\right]\,dx\,dt$$

Our goal is to minimize the objective functional, i.e.,

$$J(h^*) = \min_{h \in U} J(h)$$

with control set $U = H^1(Q)$.

This example is designed to illustrate the technique with more regular controls and to generalize the recent flow applications. The existence of an optimal control is proven. The optimality system is driven by differentiating the objective functional with respect to the control and differentiating the function mapping the control to the state. We obtain the characterization of an optimal control through an elliptic differential equation. Thus the optimality system will consist of two parabolic equations coupled with an elliptic equation.

Bibliography

[1] Kamien, M.I. and Schwarz, N. L., *Dynamic Optimization*, North-Holland, Amsterdam 1991.

[2] Lions, J. L., *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, New York, 1971.

[3] Lions, J. L., *Some aspects of the Optimal Control of Distributed Parameter Systems* , SIAM, Philadelphia, 1972.

[4] Li, X., Yong, J., *Optimal Control Theory for Infinite Dimensional Systems*, Birkhäuser, 1995.

[5] Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F., *The Mathematical Theory of Optimal Processes*, Vol 4, Gordon and Breach Science Publishers, 1986.

[6] Tikhonov, A. N. and Arsenin, V. Y., *Solutions of Ill-posed Problems*, John Wiley, New York 1997.

Part II

Optimal Control of an HIV Immunology Model

1. Introduction

AIDS - Acquired Immunity Deficiency Syndrome - is a disease that has affected the whole world in the 20 years since it was first detected. It is caused by Human Immunodeficiency Virus (HIV). Of the 34.3 million people worldwide living with HIV infection today, more than 24 million are in the developing world.

There is still a near-silence about major advances in the search for an anti-HIV vaccine, but scientists are developing new methods of treating the infected patients. Most of the chemotherapies are aimed at killing or halting the pathogen, but treatment which can boost the immune system can serve to help the body fight infection on its own [**9**]. The new treatments are aimed to reduce viral population and boost up the immune response. This brings new hope to the treatment of HIV infection, and we are exploring strategies for such type of treatments using optimal control techniques.

An ordinary differential equation (ODE) model (taken from Kirschner and Webb model [**9**]), which describes the interaction of $CD4^{+}T$ cells and HIV in the immune system is utilized, and optimal control of this ODE model is explored.

Once HIV enters the body, the human immune system tries to get rid of it. The invasion is reported to $CD4^{+}T$ cells. The $CD4$ is a "protein marker in the surface of the T cell", and the letter T refers to thymus, the organ responsible for maturing these cells after they migrate from the bone marrow (where they are manufactured). They arrive in the thymus gland where they mature. The surface of $CD4^{+}T$ possesses a protein that can bind to foreign substances such as HIV. The HIV needs a host in order to reproduce and above mentioned protein provides shelter. The HIV virus is a retrovirus, the RNA of virus is converted into DNA inside the $CD4^{+}T$ cell. Thus, when infected $CD4^{+}T$ cells begin to multiply to fight this pathogen, they produce more virus. See [**3, 7, 8, 9, 11**] for more details on disease progression.

Let T, V represent the concentration of the uninfected $CD4^{+}T$ cells and free infectious virus particles respectively, and u_1, u_2 represent two different treatment strategies. As our control classes we choose measurable functions defined on a fixed interval (as treatments can not be continued for infinite time period due to hazardous side effects) satisfying $0 \leq a_i \leq u_i(t) \leq b_i < 1$ for $i = 1, 2$. For most of HIV chemotherapy drugs, the length of treatment is less then 500 days [**3**].

The state system is

$$\frac{dT(t)}{dt} = s_1 - \frac{s_2 V(t)}{B_1 + V(t)} - \mu T(t) - kV(t)T(t) + u_1(t)T(t) \tag{1.1}$$

$$\frac{dV(t)}{dt} = \frac{g(1 - u_2(t))V(t)}{B_2 + V(t)} - cV(t)T(t) \tag{1.2}$$

satisfying $T(0) = T_0$ and, $\quad V(0) = V_0$, where T represents the concentration of $CD4^{+}T$ cells, V the concentration of HIV particles. The term $s_1 - \frac{s_2 V(t)}{B_1 + V(t)}$ is the source/proliferation of unaffected T cells , $-\mu T(t)$ is the natural loss of uninfected T cells, $-kV(t)T(t)$ is loss by infection, $\frac{gV(t)}{B_2 + v(t)}$ is viral contribution to plasma and $-cV(t)T(t)$ is the viral loss. Similarly, μ is the death rate of T cells, k is the infection rate of T cells, g is the input rate of external virus source , c is the loss rate of virus and B_1, B_2 are half saturation constants. The definitions and numerical data for the parameters are from Kirschner and Webb [**9**].

The objective functional to be maximized is

$$J(u_1, u_2) = \int_0^{t_f} [T - (A_1 u_1^2 + A_2 u_2^2)]\, dt \tag{1.3}$$

The first term represents benefit of T cells and the other terms are systemic costs of the drug treatments, where t_f is the final time. The positive constants A_1 and A_2 balance the size of the terms, and u_1^2, u_2^2 reflect the severity of the side effects of the drugs. When drugs such as interleukin (IL-2) are administered in high dose, they are toxic to the human body, which justifies the quadratic terms in the functional. Our goal is maximizing the number of T cells and minimizing the systemic cost to the body. We seek an optimal control pair,

u_1^*, u_2^* such that

$$J(u_1^*, u_2^*) = \max\{J(u_1, u_2) | (u_1, u_2) \in U\}$$

where $U = \{(u_1, u_2) | u_i$ measurable , $a_i \leq u_i \leq b_i, \quad t \in [0, t_f],$ for $i = 1, 2\}$ is the control set.

See Kirschner and Webb [**9**] for treatment strategies using immune boosting and delaying AIDS progression. They developed a mathematical model of dynamics of disease progression and IL-2 treatment of the HIV-infected immune system. Their model is based on the key markers of HIV progression, $CD4^+T$ cell level and viral levels in the plasma, which agrees with preliminary results from clinical trials. They also predict that immunotherapy administered during the early stages of disease progression is most beneficial for raising $CD4^+T$ cell count.

See references [**2**], [**3**] and [**7**] for control problems on HIV infection in different types of models using treatment with a single drug and similar objective functional. In this paper we consider two controls, one boosts the immune system and other delays the HIV progression.

Wai-Yuan and Zhinhua [**13**] have developed a stochastic model for the HIV pathogenesis under anti-viral drugs. They estimated the numbers of infectious free HIV and non-infectious free HIV as well as the number of T cells through the extended Kalman filter method. Berman [**1**] also used a stochastic model with a three-stage diffusion process (pre-treatment interval, the control treatment interval and post treatment interval) for the T cell level of an HIV-infected individual. His goal was to determine the best time to intervene with an antiviral drug using a Brownian motion representation of the T cell level.

In section 2, we investigate the existence of an optimal control pair. In section 3, we seek to maximize the objective functional and the optimal control is characterized using Pontryagin's Maximum Principle [**12**]. The uniqueness of the optimality system is proved in section 4, and some numerical results are illustrated in the last section.

2. Existence of an Optimal Control Pair

The boundedness of solutions of the system (1.1) and (1.2) for the finite time interval is used to prove the existence of an optimal control pair. This can be proved using the fact that supersolutions $\overline{T}$ and $\overline{V}$ satisfying

$$\begin{aligned} \frac{d\overline{T}}{dt} &= s_1 + u_1(t)\overline{T} \\ \frac{d\overline{V}}{dt} &= \frac{g\overline{V}}{B_2 + \overline{V}} \end{aligned}$$

are bounded on a finite time interval.

The existence of the optimal control pair can be obtained by using a result by Fleming and Rishel ([**4**, Th. 4.1, p. 68-69]).

THEOREM 2.1. *Consider the control problem with system equations (1.1),(1.2). There exists $\vec{u}^* = (u_1^*, u_2^*) \in U$ such that*

$$\max_{(u_1,u_2)\in U} J(u_1, u_2) = J(u_1^*, u_2^*).$$

Proof: In order to verify the conditions to use the result from [**4**], first, we note that the solutions are bounded. Second, we use a result by Lukes ([**10**, Th 9.2.1, p. 182] to give the existence of solutions of ODE's (1.1) and (1.2) with bounded coefficients. Note that U is bounded and convex. Since our state system is bilinear in u_1, u_2, the RHS of (1.1) and (1.2) is continuous and it can be rewritten as $\vec{f}(t, \vec{T}, \vec{u}) = \vec{\alpha}(t, \vec{T}) + \vec{\beta}(t, \vec{T})\vec{u}$, and the boundedness of solutions gives $\left|\vec{f}(t, \vec{T}, \vec{u})\right| \leq C_1 \left(1 + |\vec{T}| + |\vec{u}|\right)$, for $0 \leq t \leq t_f$, $\vec{T} \in \mathbb{R}^2$, $\vec{u} \in \mathbb{R}^2$ (where $\vec{T} = (T, V)$ and $\vec{u} = (u_1, u_2)$).

In order to verify the convexity of $\mathcal{L}$, the integrand of our objective functional, we show

$$\mathcal{L}(t, \vec{T}, (1-\varepsilon)\vec{u} + \varepsilon\vec{v}) \geq (1-\varepsilon)\mathcal{L}(t, \vec{T}, \vec{u}) + \varepsilon\mathcal{L}(t, \vec{T}, \vec{v}) \quad \text{for } 0 < \varepsilon < 1 \tag{2.1}$$

where $\mathcal{L}(t, \vec{T}, \vec{u}) = [T - (A_1 u_1^2 + A_2 u_2^2)]$.

$$\mathcal{L}(t, \vec{T}, (1-\varepsilon)\vec{u} + \varepsilon\vec{v})$$

$$= \left[T - A_1((1-\varepsilon)u_1 + \varepsilon v_1)^2 - A_2((1-\varepsilon)u_2 + \varepsilon v_2)^2\right]$$

$$= T - (A_1 u_1^2 + A_2 u_2^2) - \Big(A_1[(\varepsilon^2 - 2\varepsilon)u_1^2 + \varepsilon^2 v_1^2 + 2\varepsilon(1-\varepsilon)u_1 v_1]$$

$$+A_2[(\varepsilon^2 - 2\varepsilon)u_2^2 + \varepsilon^2 v_2^2 + 2\varepsilon(1-\varepsilon)u_2 v_2]\Big)$$

and

$$\begin{aligned} & (1-\varepsilon)\mathcal{L}(t,\vec{T},\vec{u}) + \varepsilon\mathcal{L}(t,\vec{T},\vec{v}) \\ = \;\; & (1-\varepsilon)[T - (A_1 u_1^2 + A_2 u_2^2)] + \varepsilon[T - (A_1 v_1^2 + A_2 v_2^2)] \\ = \;\; & T - (A_1 u_1^2 + A_2 u_2^2) - \varepsilon(-A_1 u_1^2 - A_2 u_2^2 + A_1 v_1^2 + A_2 v_2^2). \end{aligned}$$

Thus in order to show that $\mathcal{L}(t,\vec{T},\cdot)$ is convex in U, we note that following inequality holds:

$$A_1[(\varepsilon^2 - 2\varepsilon)u_1^2 + \varepsilon^2 v_1^2 + 2\varepsilon(1-\varepsilon)u_1 v_1] + A_2[(\varepsilon^2 - 2\varepsilon)u_2^2 + \varepsilon^2 v_2^2 + 2\varepsilon(1-\varepsilon)u_2 v_2]$$
$$\leq \varepsilon(-A_1 u_1^2 - A_2 u_2^2 + A_1 v_1^2 + A_2 v_2^2).$$

In other words, we need to show

$$A_1(\varepsilon^2 - \varepsilon)(u_1^2 + v_1^2) + A_2(\varepsilon^2 - \varepsilon)(u_2^2 + v_2^2 + 2\varepsilon(1-\varepsilon)(A_1 u_1 v_1 + A_2 u_2 v_2) \leq 0$$

which is equivalent to

$$-A_1\left(\sqrt{\varepsilon(1-\varepsilon)}u_1 - \sqrt{\varepsilon(1-\varepsilon)}v_1\right)^2 - A_2\left(\sqrt{\varepsilon(1-\varepsilon)}u_2 - \sqrt{\varepsilon(1-\varepsilon)}v_2\right)^2 \leq 0.$$

The above inequality holds since $A_1, A_2 > 0$. Hence (2.1) holds.

Finally, we need to show that $\mathcal{L}(t,\vec{T},\vec{u}) \leq c_2 - c_1|\vec{u}|^\beta$, where $c_1 > 0$ and $\beta > 1$. For our case:

$$\mathcal{L}(t,\vec{T},\vec{u}) \;\; = \;\; T - (A_1 u_1^2 + A_2 u_2^2) \leq c_2 - c_1|\vec{u}|^2$$

where c_2 depends on the upper bound on T, and $c_1 > 0$ since $A_1, A_2 > 0$. We conclude there exists an optimal control pair by the existence result from [**4**, Theorem 4.1]. $\square$

3. Optimality System

In section 2, we proved the existence of an optimal control pair system for maximizing the functional (1.3) subject to (1.1)-(1.2). In order to derive the necessary conditions on this optimal control pair, we use Pontryagin's Maximum Principle [**12**].

The Lagrangian is defined as following:

$$\begin{aligned} L &= \left[T - (A_1 u_1^2 + A_2 u_2^2)\right] \\ &+ \lambda_1 \left[s_1 - \frac{s_2 V(t)}{B_1 + V(t)} - \mu T(t) - kV(t)T(t) + u_1(t)T(t)\right] \\ &+ \lambda_2 \left[\frac{g(1-u_2(t))V(t)}{B_2 + V(t)} - cV(t)T(t)\right] \\ &+ w_{11}(t)(b_1 - u_1) + w_{12}(t)(u_1 - a_1) \\ &+ w_{21}(t)(b_2 - u_2) + w_{22}(t)(u_2 - a_2), \end{aligned}$$

where $w_{11}(t), w_{12}(t), w_{21}(t), w_{22}(t) \geq 0$ are penalty multipliers satisfying

$$w_{11}(t)(b_1 - u_1) = 0, w_{12}(t)(u_1 - a_1) = 0 \quad \text{at } u_1^*$$

and

$$w_{21}(t)(b_2 - u_2) = 0, w_{22}(t)(u_2 - a_2) = 0 \quad \text{at } u_2^* \ [11].$$

THEOREM 3.1. *Given optimal controls u_1^*, u_2^* and solutions T^*, V^* of the corresponding state system (1.1)-(1.2), there exist adjoint variables λ_1, λ_2 satisfying*

$$\begin{aligned} \lambda_1' &= -1 + \lambda_1 \left[\mu + kV^*(t) - u_1^*(t)\right] + \lambda_2 c V^*(t) \\ \lambda_2' &= \lambda_1 \left[\frac{B_1 s_2}{(b_1 + V^*(t))^2} + kT^*(t)\right] - \lambda_2 \left[\frac{B_2 g(1-u_2^*(t))}{(B_2 + V^*(t))^2} - cT^*(t)\right] \end{aligned}$$

and $\lambda_1(t_f) = \lambda_2(t_f) = 0$, the transversality conditions. Furthermore

$$u_1^*(t) = \min\left\{\max\left\{a_1, \frac{1}{2A_1}\left(\lambda_1 T^*(t)\right)\right\}, b_1\right\}$$

$$u_2^*(t) = \min\left\{\max\left\{a_2, -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{(B_2+V^*(t))}\right\}, b_2\right\}$$

Proof: The form of the adjoint equations and transvarsality conditions are standard results from Pontryagin's Maximum Principle [**12**]. We differentiate the Lagrangian with respect to states, T and V respectively.

$$\begin{aligned}
\frac{\partial L}{\partial T} &= 1 + \lambda_1\left[-\mu - kV(t) + u_1(t)\right] - \lambda_2 cV(t) \\
\frac{\partial L}{\partial V} &= -\lambda_1\left[\frac{(B_1+V(t))s_2 - s_2V(t)}{(B_1+v(t))^2} + kT(t)\right] \\
&\quad +\lambda_2\left[\frac{(B_2+V(t))g(1-u_2(t)) - g(1-u_2(t))V(t)}{(B_2+V(t))^2} - cT(t)\right] \\
&= -\lambda_1\left[\frac{B_1s_2}{(B_1+V(t))^2} + kT(t)\right] + \lambda_2\left[\frac{B_2g(1-u_2(t))}{(B_2+V(t))^2} - cT(t)\right]
\end{aligned}$$

Thus the adjoint system can be written as:

$$\lambda_1' = -\frac{\partial L}{\partial T} = -1 + \lambda_1\left[\mu + kV^*(t) - u_1(t)\right] + \lambda_2 cV^*(t) \tag{3.1}$$

$$\lambda_2' = -\frac{\partial L}{\partial V} = \lambda_1\left[\frac{B_1s_2}{(b_1+V^*(t))^2} + kT^*(t)\right] - \lambda_2\left[\frac{B_2g(1-u_2(t))}{(B_2+V^*(t))^2} - cT^*(t)\right] \tag{3.2}$$

The optimality equations [**6**] are :

$$\begin{aligned}
\frac{\partial L}{\partial u_1} &= -2A_1u_1(t) + \lambda_1T^*(t) - w_{11}(t) + w_{12}(t) = 0 \qquad \text{at } u_1^* \\
\frac{\partial L}{\partial u_2} &= -2A_2u_2(t) + \lambda_2\left[-\frac{gV^*(t)}{B_2+V^*(t)}\right] - w_{21}(t) + w_{22}(t) = 0 \quad \text{at } u_2^*.
\end{aligned}$$

Hence we obtain

$$u_1^*(t) = \frac{1}{2A_1}\left[\lambda_1T^*(t) - w_{11}(t) + w_{12}(t)\right] \tag{3.3}$$

$$u_2^*(t) = \frac{1}{2A_2}\left[-\lambda_2\frac{gV^*(t)}{B_2+V^*(t)} - w_{21}(t) + w_{22}(t)\right]. \tag{3.4}$$

There are three cases for u_1^* at any time t:

Case(i): $a_1 = u_1^*(t)$.

Since $u_1^*(t) \neq b_1$, $w_{11}(t) = 0$. Then $a_1 = u_1^*(t) = \dfrac{1}{2A_1}\left[\lambda_1 T^*(t) + w_{12}(t)\right]$ by (3.3). Then solving for $w_{12}(t)$ gives

$$2A_1 a_1 - \lambda_1 T^*(t) = w_{12}(t) \geq 0,$$

which implies

$$2A_1 a_1 \geq \lambda_1 T^*(t)$$

and

$$a_1 \geq \frac{1}{2A_1}\lambda_1 T^*(t).$$

Case(ii): $a_1 < u_1^*(t) < b_1$.

By the definition of the penalty multipliers, $w_{12}(t) = w_{11}(t) = 0$, and (3.3) gives

$$u_1^* = \frac{1}{2A_1}\lambda_1 T^*(t)$$

Case(iii): $u_1^*(t) = b_1$

Since $u_1^*(t) \neq a_1$, $w_{12}(t) = 0$, and (3.3) gives

$$b_1 = u_1^*(t) = \frac{1}{2A_1}\left[\lambda_1 T^*(t) - w_{11}(t)\right]$$

which implies

$$0 \leq w_{11} = \lambda_1 T^*(t) - 2A_1 b_1$$

and

$$b_1 \leq \frac{\lambda_1 T^*(t)}{2A_1}.$$

Hence we conclude

$$u_1^* = \begin{cases} \frac{1}{2A_1}\lambda_1 T^*(t) & \text{if} \quad a_1 < \frac{1}{2A_1}\lambda_1 T^*(t) < b_1 \\ a_1 & \text{if} \quad \frac{1}{2A_1}\lambda_1 T^*(t) \leq a_1 \\ b_1 & \text{if} \quad \frac{1}{2A_1}\lambda_1 T^*(t) \geq b_1. \end{cases}$$

In compact notation, $u_1^*(t) = \min\left\{\max\left\{a_1, \frac{1}{2A_1}\left(\lambda_1 T^*(t)\right)\right\}, b_1\right\}$.

Similar to the $u_1^*(t)$, there are three cases for $u_2^*(t)$:

Case(i): $a_2 = u_2^*(t)$.

We have $w_{21}(t) = 0$, and then $a_2 = u_2^*(t) = \frac{1}{2A_2}\left[-\lambda_2 \frac{V^*(t)}{B_2 + V^*(t)} + w_{22}(t)\right]$ by (3.4). Then solving for $w_{22}(t)$ gives

$$2A_2 a_2 + \lambda_2 \frac{V^*(t)}{B_2 + V^*(t)} = w_{22}(t) \geq 0,$$

which implies

$$2A_2 a_2 \geq -\lambda_2 \frac{V^*(t)}{B_2 + V^*(t)}$$

and

$$a_2 \geq -\frac{1}{2A_2}\lambda_2 \frac{V^*(t)}{B_2 + V^*(t)}.$$

Case(ii): $a_2 < u_2^*(t) < b_2$.

We have $w_{21}(t) = w_{22}(t) = 0$ and

$$u_2^* = -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{B_2 + V^*(t)}.$$

Case(iii): $u_2^*(t) = b_2$

Since $u_2^*(t) \neq a_2$, $w_{22}(t) = 0$. Characterization (3.3) gives

$$b_2 = u_2^*(t) = \frac{1}{2A_2}\left[\lambda_2 \frac{V^*(t)}{B_2 + V^*(t)} - w_{21}(t)\right]$$

which implies

$$0 \leq w_{21} = \lambda_2 \frac{V^*(t)}{B_2 + V^*(t)} - 2A_2 b_2 \text{ (since } w_{ij} \geq 0,\ \ i, j = 1, 2)$$

and

$$b_2 \leq -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{B_2 + V^*(t)}$$

Hence, we conclude

$$u_2^* = \begin{cases} -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{B_2+V^*(t)} & \text{if} \quad a_2 < -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{B_2+V^*(t)} < b_2 \\ a_2 & \text{if} \quad -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{B_2+V^*(t)} \leq a_2 \\ b_2 & \text{if} \quad -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{B_2+V^*(t)} \geq b_2. \end{cases}$$

In compact notation, $u_2^*(t) = \min\left\{\max\left\{a_2, -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{B_2+V^*(t)}\right\}, b_2\right\}$. □

The optimality system consists of the state system coupled with the adjoint system with the initial and transversality conditions together with the characterization of the optimal control pair

$$u_1^*(t) = \min\left\{\max\left\{a_1, \frac{1}{2A_1}(\lambda_1 T^*(t))\right\}, b_1\right\} \tag{3.5}$$

$$u_2^*(t) = \min\left\{\max\left\{a_2, -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{B_2+V^*(t)}\right\}, b_2\right\}. \tag{3.6}$$

Utilizing (3.5) and (3.6), we have the following optimality system which characterizes the optimal control.

$$\begin{aligned} \frac{dT(t)}{dt} &= s_1 - \frac{s_2 V(t)}{B_1+V(t)} - \mu T(t) - kV(t)T(t) \\ &\quad + \min\left\{\max\left\{a_1, \frac{1}{2A_1}(\lambda_1 T^*(t))\right\}, b_1\right\} T(t) \\ \frac{dV(t)}{dt} &= \frac{g\left(1-\min\left\{\max\left\{a_2, -\frac{\lambda_2}{2A_2}\frac{V^*(t)}{B_2+V^*(t)}\right\}, b_2\right\}\right)V(t)}{B_2+V(t)} - cV(t)T(t) \\ \lambda_1' &= -1 + \lambda_1\left[\mu + kV(t) - \min\left\{\max\left\{a_1, \frac{1}{2A_1}(\lambda_1 T^*(t))\right\}, b_1\right\}\right] \\ &\quad + \lambda_2 cV(t) \qquad\qquad (3.7) \\ \lambda_2' &= \lambda_1\left[\frac{B_1 s_2}{(B_1+V(t))^2} + kT(t)\right] \\ &\quad -\lambda_2\left[\frac{B_2 g\left(1-\min\left\{\max\left\{a_2, -\frac{\lambda_2}{2A_2}\frac{V(t)}{B_2+V(t)}\right\}, b_2\right\}\right)}{(B_2+V(t))^2} - cT(t)\right] \\ \lambda_1(t_f) &= \lambda_1(t_f) = 0, \text{ and } T(0) = T_0, V(0) = V_0. \end{aligned}$$

By therorems 2.1 and 2.1 the solution of optimality system (3.7) exists.

4. Uniqueness of the Optimality System

We will use the following lemma while proving the uniqueness of solutions of the optimality system for the small time interval.

LEMMA 4.1. *The function* $u^*(s) = \min(\max(s, a), b))$ *is Lipschitz continuous in* s*, where* $a < b$ *are some fixed positive constants.*

Proof: Consider s_1, s_2 as real numbers and a, b as fixed positive constants. We will show the Lipschitz continuity holds in all possible cases for $\max(s, a)$. Similar arguments hold for $\min(\max(s, a), b))$ as well.

1: $s_1 \geq a,\ s_2 \geq a$.

$|\max(s_1, a) - \max(s_2, a)| = |s_1 - s_2|$

2: $s_1 \geq a,\ s_2 \leq a$.

$|\max(s_1, a) - \max(s_2, a)| = |s_1 - a| \leq |s_1 - s_2|$

3: $s_1 \leq a,\ s_2 \geq a$.

$|\max(s_1, a) - \max(s_2, a)| = |a - s_2| \leq |s_1 - s_2|$

4: $s_1 \leq a,\ s_2 \leq a$.

$|\max(s_1, a) - \max(s_2, a)| = |a - a| = 0 \leq |s_1 - s_2|$

Hence $|\max(s_1, a) - \max(s_2, a)| \leq |s_1 - s_2|$ and we have Lipschitz continuity of u^* in s. □

THEOREM 4.1. *For* t_f *sufficiently small, bounded solutions to the optimality system are unique.*

Proof: Suppose $(T, V, \lambda_1, \lambda_2)$ and $(\overline{T}, \overline{V}, \overline{\lambda}_1, \overline{\lambda}_2)$ are two different solutions of our optimality system (3.7). Let $T = e^{\lambda t}p, V = e^{\lambda t}q, \lambda_1 = e^{-\lambda t}w, \lambda_2 = e^{-\lambda t}z$ and $\overline{T} = e^{\lambda t}\overline{p}, \overline{V} =$

$e^{\lambda t}\overline{q}, \overline{\lambda}_1 = e^{-\lambda t}\overline{w}, \overline{\lambda}_2 = e^{-\lambda t}\overline{z}$, where $\lambda > 0$ is to be chosen. Further we let

$$\begin{aligned}
u_1^*(t) &= \min\left\{\max\left\{a_1, \frac{1}{2A_1}(wp)\right\}, b_1\right\} \\
u_2^*(t) &= \min\left\{\max\left\{a_2, -\frac{z}{2A_2}\frac{q}{B_2 + e^{\lambda t}q}\right\}, b_2\right\}
\end{aligned}$$

and

$$\begin{aligned}
\overline{u}_1^*(t) &= \min\left\{\max\left\{a_1, \frac{1}{2A_1}(\overline{w}\,\overline{p})\right\}, b_1\right\} \\
\overline{u}_2^*(t) &= \min\left\{\max\left\{a_2, -\frac{\overline{z}}{2A_2}\frac{\overline{q}}{B_2 + e^{\lambda t}\overline{q}}\right\}, b_2\right\}.
\end{aligned}$$

Now we substitute $T = e^{\lambda t}p$ into the first ODE of (3.7) and get

$$\begin{aligned}
e^{\lambda t}(p' + \lambda p) &= s_1 - \frac{s_2 e^{\lambda t} q}{B_1 + e^{\lambda t} q} - \mu e^{\lambda t} p - k e^{2\lambda t} pq \\
&\quad + \min\left\{\max\left\{a_1, \frac{1}{2A_1}\left(e^{-\lambda t} w e^{\lambda t} p\right)\right\}, b_1\right\} e^{\lambda t} p
\end{aligned}$$

or

$$\begin{aligned}
p' + \lambda p &= e^{-\lambda t} s_1 - \frac{s_2 q}{B_1 + e^{\lambda t} q} - \mu p - k e^{\lambda t} pq \\
&\quad + \min\left\{\max\left\{a_1, \frac{1}{2A_1}(wp)\right\}, b_1\right\} p. \qquad (4.1)
\end{aligned}$$

Similarly, for $V = e^{\lambda t}q, \lambda_1 = e^{-\lambda t}w, \lambda_2 = e^{-\lambda t}z$, we obtain

$$\begin{aligned}
q' + \lambda q &= \frac{g\left(1 - \min\left\{\max\left\{a_2, -\frac{z}{2A_2}\frac{q}{B_2+e^{\lambda t}q}\right\}, b_2\right\}\right) q}{B_2 + e^{\lambda t} q} - c e^{\lambda t} pq \qquad (4.2) \\
-w' + \lambda w &= e^{\lambda t} - w\left[\mu + k e^{\lambda t} q - \min\left\{\max\left\{a_1, \frac{1}{2A_1}(pw)\right\}, b_1\right\}\right] \\
&\quad - c e^{\lambda t} q z \qquad (4.3) \\
-z' + \lambda z &= -w\left[\frac{B_1 s_2}{(B_1 + e^{\lambda t} q)^2} + k e^{\lambda t} p\right] \\
&\quad + z\left[\frac{B_2 g\left(1 - \min\left\{\max\left\{a_2, -\frac{z}{2A_2}\frac{q}{B_2+e^{\lambda t}q}\right\}, b_2\right\}\right)}{(B_2 + e^{\lambda t} q)^2} - c e^{\lambda t} p\right] \qquad (4.4)
\end{aligned}$$

Now we subtract the equations for T and $\overline{T}$, V and $\overline{V}$, λ_1 and $\overline{\lambda}_1$, λ_2 and $\overline{\lambda}_2$. Then multiply each equation by an appropriate difference of functions and integrate from 0 to t_f. Next we add all four integral equations and will use estimates to obtain uniqueness.

Using Lemma 4.1, we have

$$|u_1^*(t) - \overline{u}_1^*(t)| \leq \frac{1}{2A_1} |wp - \overline{wp}|$$

and

$$\begin{aligned} |u_2^*(t) - \overline{u}_2^*(t)| &\leq \left| \frac{1}{2A_2} \left(\frac{zq}{B_2 + e^{\lambda t} q} - \frac{\overline{zq}}{B_2 + e^{\lambda t}\overline{q}} \right) \right| \\ &\leq \frac{1}{2A_2} \left| \frac{B_2(zq - \overline{zq}) + e^{\lambda t} q\overline{q}(z - \overline{z})}{(B_2 + e^{\lambda t} q)\,(B_2 + e^{\lambda t}\overline{q})} \right|. \end{aligned}$$

Then we obtain

$$\begin{aligned} &\frac{1}{2}(p - \overline{p})^2(t_f) + \lambda \int_0^{t_f} (p - \overline{p})^2 \, dt \\ &\leq \int_0^{t_f} \left| \frac{s_2 q}{B_1 + e^{\lambda t} q} - \frac{s_2 \overline{q}}{B_1 + e^{\lambda t}\overline{q}} \right| |p - \overline{p}| \, dt + \int_0^{t_f} \mu |p - \overline{p}|^2 \, dt \\ &+ k \int_0^{t_f} e^{\lambda t} |pq - \overline{pq}||p - \overline{p}| \, dt + \int_0^{t_f} |u_1^* p - \overline{u_1^* p}||p - \overline{p}| \, dt \\ &\leq C_1 \int_0^{t_f} \left[|p - \overline{p}|^2 + |q - \overline{q}|^2 + |w - \overline{w}|^2 \right] \, dt \\ &+ C_2 e^{\lambda t_f} \int_0^{t_f} \left[|p - \overline{p}|^2 + |q - \overline{q}|^2 \right] \, dt, \end{aligned}$$

$$\begin{aligned} &\frac{1}{2}(q - \overline{q})^2(t_f) + \lambda \int_0^{t_f} (q - \overline{q})^2 \, dt \\ &\leq \int_0^{t_f} g \left| \frac{(1 - u_2^*(t))q}{B_2 + e^{\lambda t} q} - \frac{(1 - \overline{u}_2^*(t))\overline{q}}{B_2 + e^{\lambda t}\overline{q}} \right| |q - \overline{q}| \, dt + c \int_0^{t_f} e^{\lambda t} |pq - \overline{pq}||q - \overline{q}| \, dt \\ &\leq C_3 \int_0^{t_f} \left[|q - \overline{q}|^2 + |z - \overline{z}|^2 \right] \, dt + C_4 e^{2\lambda t_f} \int_0^{t_f} \left[|p - \overline{p}|^2 + |q - \overline{q}|^2 + |z - \overline{z}|^2 \right] \, dt, \end{aligned}$$

$$\frac{1}{2}(w - \overline{w})^2(0) + \lambda \int_0^{t_f} (w - \overline{w})^2 \, dt$$

$$\leq \mu \int_0^{t_f} |w-\overline{w}|^2\,dt + k\int_0^{t_f} e^{\lambda t}|wq-\overline{wq}||w-\overline{w}|\,dt$$
$$+\int_0^{t_f} |u_1^*-\overline{u_1^*}||w-\overline{w}|\,dt + c\int_0^{t_f} e^{\lambda t}|zq-\overline{zq}||w-\overline{w}|\,dt$$
$$\leq C_5\int_0^{t_f}\left[|p-\overline{p}|^2+|w-\overline{w}|^2\right]\,dt + C_6 e^{\lambda t_f}\int_0^{t_f}\left[|q-\overline{q}|^2+|w-\overline{w}|^2+|z-\overline{z}|^2\right]\,dt,$$

and

$$\frac{1}{2}(z-\overline{z})^2(0)+\lambda\int_0^t (z-\overline{z})^2\,dt$$
$$\leq \int_0^{t_f}\left|\frac{B_1 s_2 w}{(B_1+e^{\lambda t}q)^2}-\frac{B_1 s_2\overline{w}}{(B_1+e^{\lambda t}\overline{q})^2}\right||z-\overline{z}|\,dt + k\int_0^{t_f} e^{\lambda t}|pw-\overline{pw}||z-\overline{z}|\,dt$$
$$+B_2 g\int_0^{t_f}\left|\frac{(1-u_2^*(t))z}{(B_2+e^{\lambda t}q)^2}-\frac{(1-\overline{u}_2^*(t))\overline{z}}{(B_2+e^{\lambda t}\overline{z})^2}\right||z-\overline{z}|\,dt + c\int_0^{t_f} e^{\lambda t}|pz-\overline{pz}||z-\overline{z}|\,dt$$
$$\leq C_7\int_0^{t_f}\left[|q-\overline{q}|^2+|w-\overline{w}|^2+|z-\overline{z}|^2\right]\,dt$$
$$+C_8 e^{3\lambda t_f}\int_0^{t_f}\left[|p-\overline{p}|^2+|q-\overline{q}|^2+|w-\overline{w}|^2+|z-\overline{z}|^2\right]\,dt.$$

Adding all above four estimates gives

$$\frac{1}{2}(p-\overline{p})^2(t_f)+\frac{1}{2}(q-\overline{q})^2(t_f)+\frac{1}{2}(w-\overline{w})^2(0)+\frac{1}{2}(z-\overline{z})^2(0)$$
$$+\lambda\int_0^{t_f}\left[(p-\overline{p})^2+(q-\overline{q})^2+(w-\overline{w})^2+(z-\overline{z})^2\right]\,dt$$
$$\leq \left(\tilde{C}_1+\tilde{C}_2 e^{3\lambda t_f}\right)\int_0^{t_f}\left[(p-\overline{p})^2+(q-\overline{q})^2+(w-\overline{w})^2+(z-\overline{z})^2\right]\,dt.$$

Thus from above equation we conclude that

$$\left(\lambda-\tilde{C}_1-\tilde{C}_2 e^{3\lambda t_f}\right)\int_0^{t_f}\left[(p-\overline{p})^2+(q-\overline{q})^2+(w-\overline{w})^2+(z-\overline{z})^2\right]\,dt\leq 0$$

where $\tilde{C}_1, \tilde{C}_2$ depend on the coefficients and the bounds on p, q, w, z.

If we choose λ such that $\lambda > \tilde{C}_1+\tilde{C}_2$ and $t_f < \frac{1}{3\lambda}\ln\left(\frac{\lambda-\tilde{C}_1}{\tilde{C}_2}\right)$, then $p=\overline{p}$, $q=\overline{q},\ w=\overline{w},\ z=\overline{z}$. Hence the solution is unique for small time. $\square$

See Fister *et.al* [**3**] for proof of a similar uniqueness result. The uniqueness for a small time interval is usual in “two-point” boundary value problems due to opposite time orientations; the state equations have initial conditions and the adjoint equations have final time conditions. The optimal controls, u_1^*, u_2^* are characterized in terms of the unique solution of the optimality system. The above optimal controls give an optimal treatment strategy for the HIV infected patient under the scenario of these two types of drug treatments.

5. Numerical Illustration

The optimality system is solved using an iterative method with a Runge-Kutta fourth order scheme. The state system with an initial guess is solved forward in time and then the adjoint system is solved backward in time. The controls are updated at the end of each iteration using the formula for optimal controls, which we derived in section 3 (in the last part of Theorem 3.1). The iterations continue until convergence is achieved. See [**5**] for background on such iterative algorithms.

Using different combinations of weight factors (A_1, A_2) and upperbounds (b_1, b_2) for controls, one can generate several treatment schedules for various time periods. Here we illustrate a case for two different values of A_1 for a 50-day treatment schedule. The other parameters in the model are [**9**]: $s_1 = 2.0, s_2 = 1.5, \mu = 0.002, k = 2.5E - 4, c = 0.007, g = 30.0, a_1 = 0.0, a_2 = 0.0, B_1 = 14.0, B_2 = 1.0$. Figures 1-4 are plotted using $A_1 = 250000, A_2 = 75, b_1 = 0.02, b_2 = 0.9$, Figures 5-8 are plotted using $A_1 = 500000$ and keeping the rest of the parameters unchanged. Figure 1 and 2 represent the controls u_1^*, u_2^* for drug administration schedule for the first set of parameters. The immune boosting drug is administered in full scale nearly up to 40 days and then is tapered off. Similarly the viral suppression drug is administered in full scale nearly up to 5 days and then is tapered off. Figure 3 represents number of T cells during our treatment period. The T cell population increases almost linearly up to 45 days and levels off after that time.

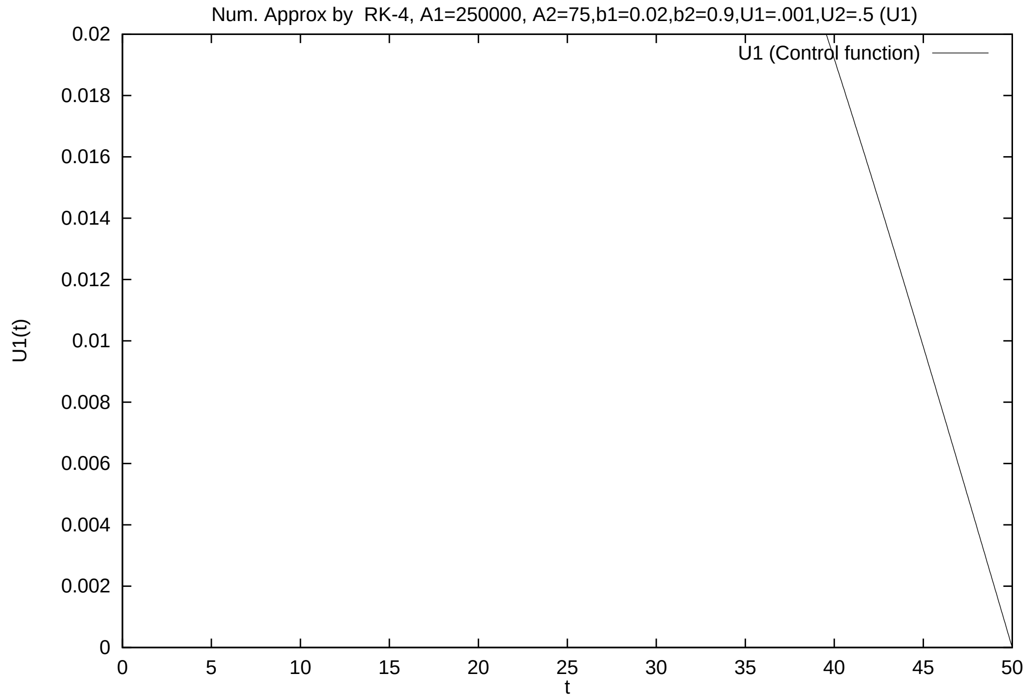

FIGURE 1. u_1 *(first control) for* $A_1 = 250000$

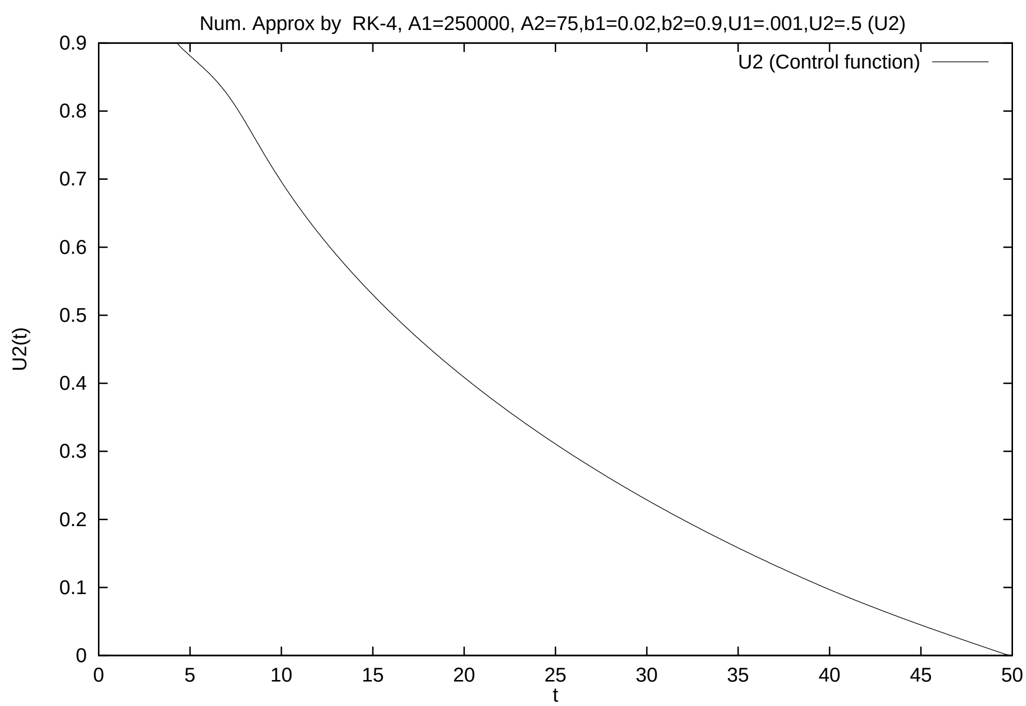

FIGURE 2. u_2 *(second control) for* $A_1 = 250000$

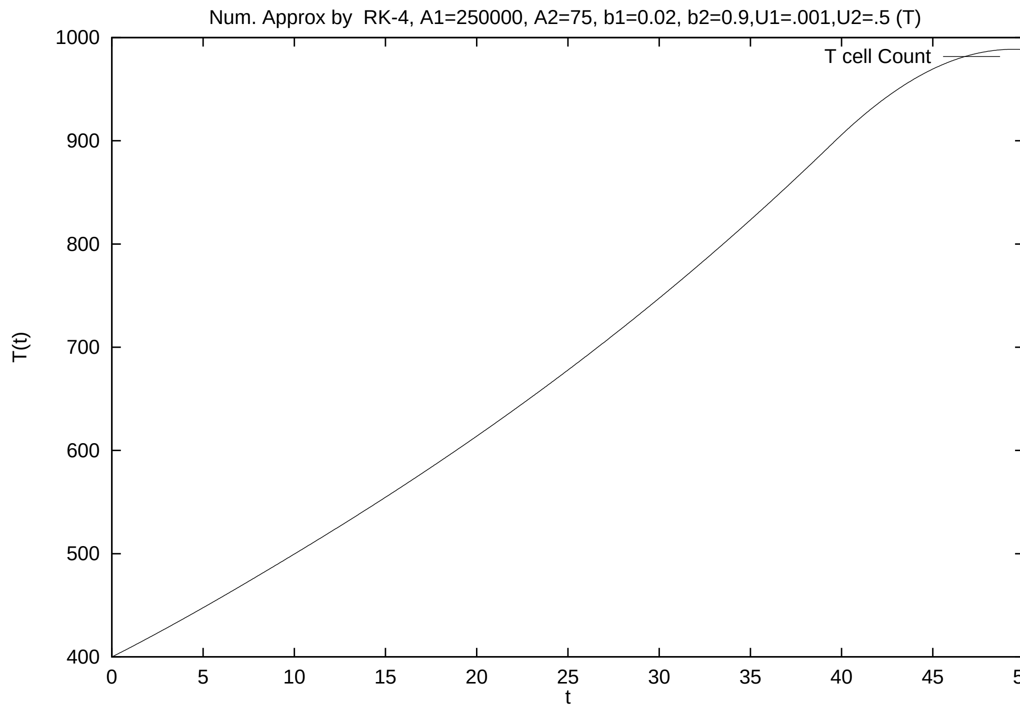

FIGURE 3. T *cell count for* $A_1 = 250000$

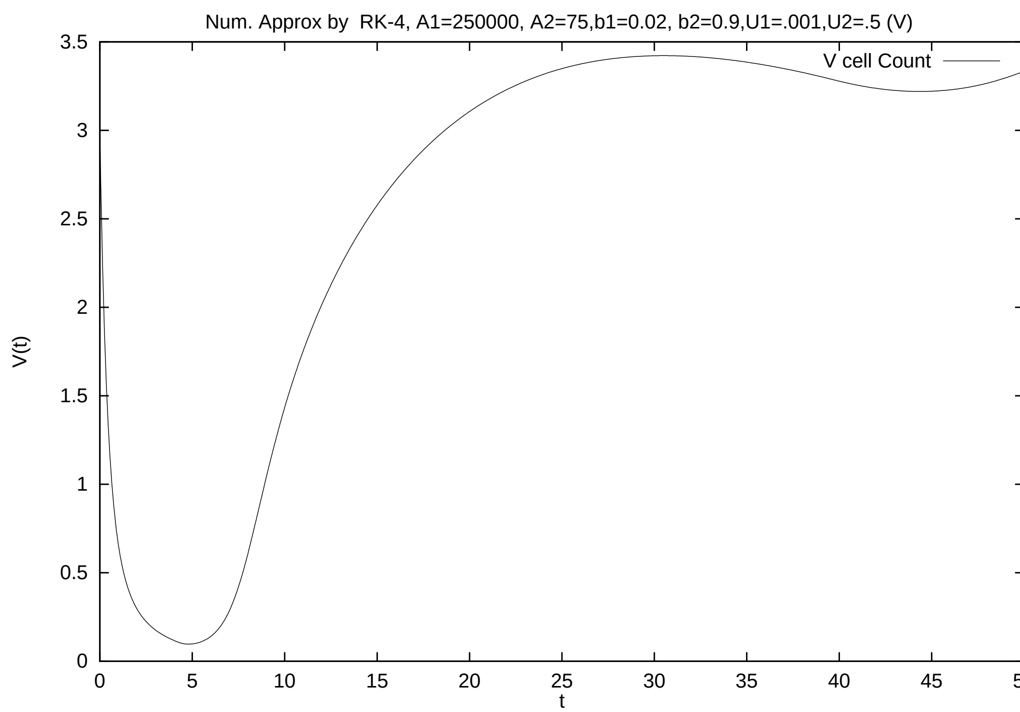

FIGURE 4. V *cell count for* $A_1 = 250000$

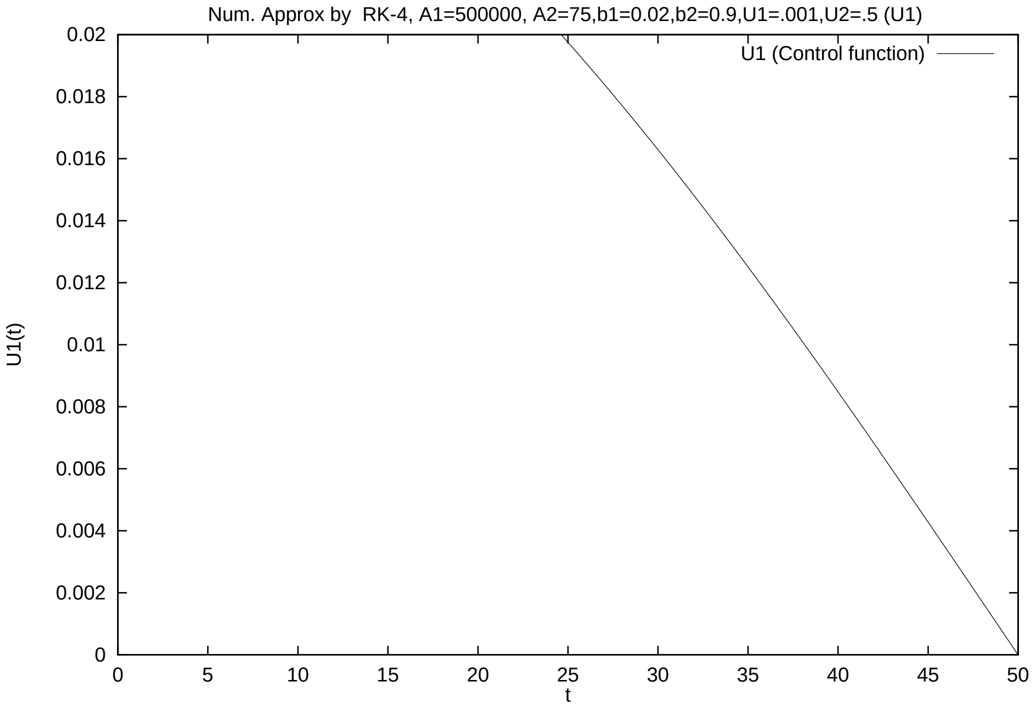

FIGURE 5. u_1 *(first control) for* $A_1 = 500000$

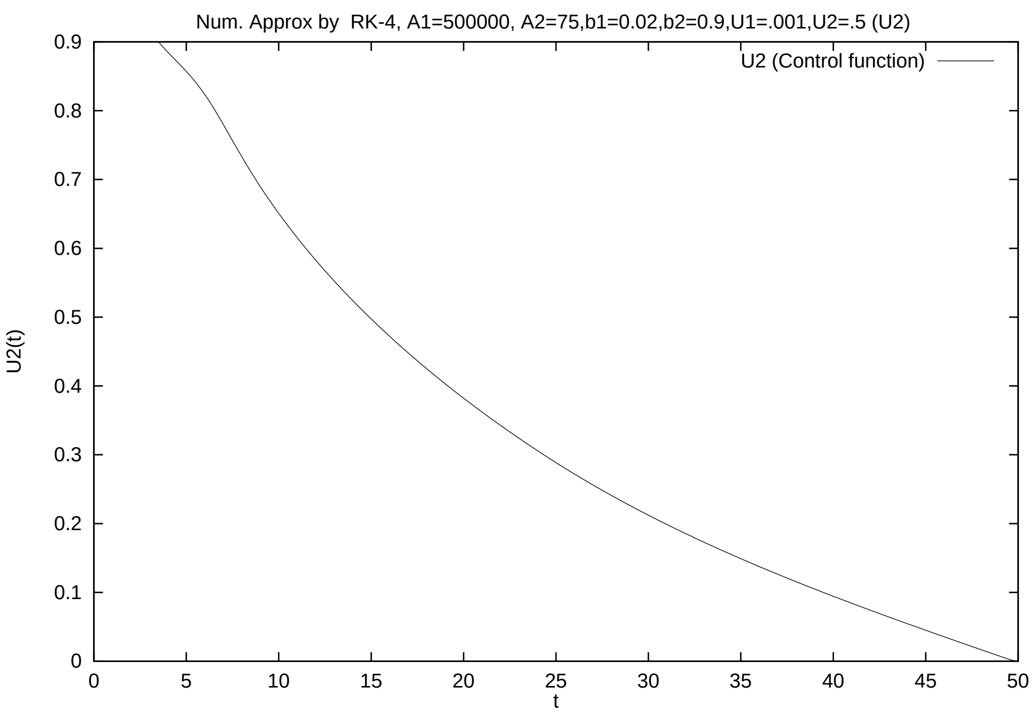

FIGURE 6. u_2 *(second control) for* $A_1 = 500000$

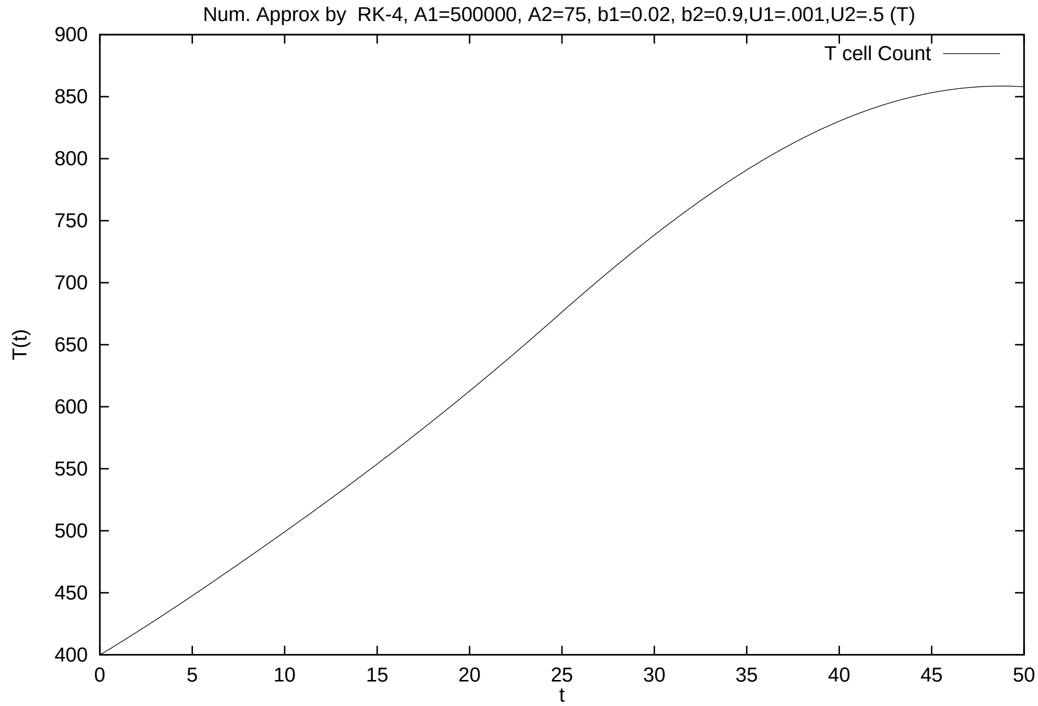

FIGURE 7. T *cell count for* $A_1 = 500000$

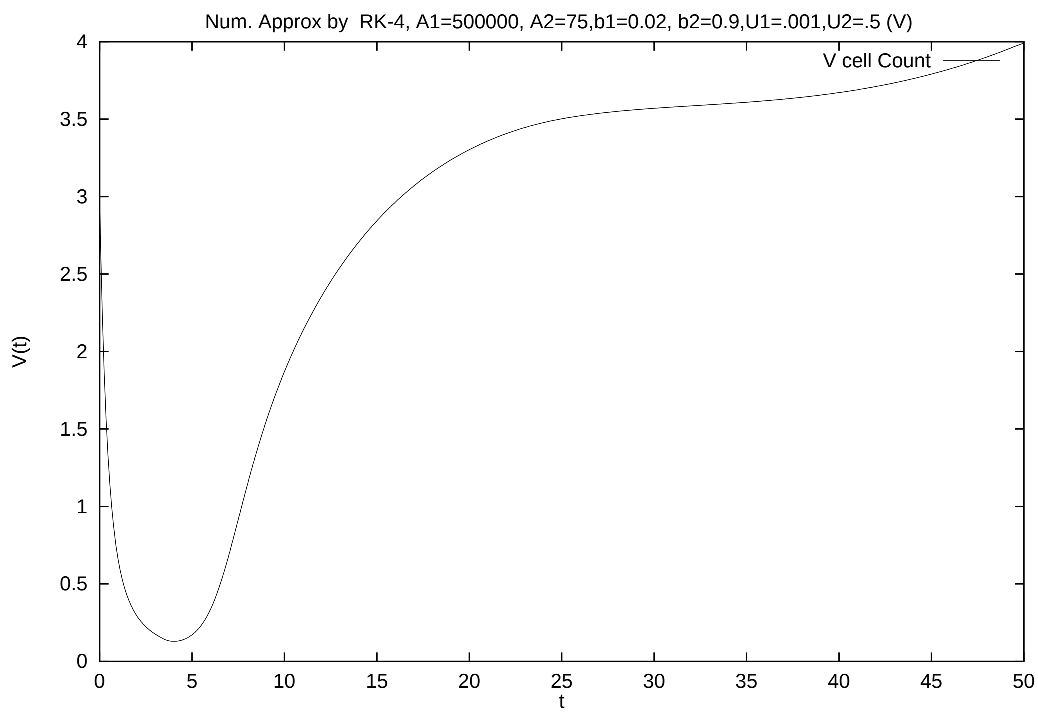

FIGURE 8. V *cell count for* $A_1 = 500000$

Figure 4 represents the virus population during our treatment period. In the beginning, we see a sharp decrease in the virus population and after few days it starts to build up again. It fluctuate some but finally it has an increasing tendency at the end of the treatment. Figure 5 and Figure 6 represent the optimal controls u_1^*, u_2^* for drug administration schedule for the second set of parameters. The immune boosting drug is administered in full scale nearly up to 25 days and then tapers off. Similarly the viral suppression drug is administered in full scale nearly up to 4 days and then tapers off. When we compare Figures 7 and 8 for T and V with Figures 3 and 4, we see that larger A_1 values reduce the T cell population and increase the virus population.

After observing results of various combinations of parameters, we conclude that the larger the weight factor, the earlier we start tapering off the treatment. This larger weight factor means that the drug is more toxic and that drug is used less. Both drug treatments are given over the same interval but the time of giving maximum level varies. The virus population drops for the first few days and starts to build up again. We can keep the virus population under a certain level but cannot eradicate it.

Bibliography

[1] Berman, S. M., *Optimal Timing of Antiviral Therapy in HIV Infection*, J. Appl. Prob. , 31A (1994), 3-15.

[2] Butler, S., Kirschner, D. and Lenhart, S., *Optimal Control of the Chemotherapy Affecting the Infectivity of HIV*, Mathematical Biology and Medicine, Vol. 6, World Scientific, 1995, 557-569.

[3] Fister, K. R., Lenhart, S. and Mc Nally, J.S., *Optimizing Chemotherapy in an HIV Model*, Electron.J. Diff. Eqns. Vol. 1998(1998), No. 32, 1-12.

[4] Fleming, W. H. and Rishel, R. W., *Deterministic and Stochastic Optimal Control*, Springer Verlag, New York, 1975.

[5] Hackbusch, W. K.,*A Numerical Method for Solving Parabolic Equations with Opposite Orientations*, Computing 20 (1978), 229-240.

[6] Kamien, M.I. and Schwarz, N. L., *Dynamic Optimization*, North-Holland, Amsterdam 1991.

[7] Kirschner, D., Lenhart, S. and Serbin,S., *Optimal Control of the Chemotherapy of HIV* J. Math. Biol (1997) 35: 775-792.

[8] Kirschner, D., Perelson, A., and Deboer R., *The Dynamics of HIV infection of* $CD4^{+}T$ *Cells*, Mathematical Biosciences **114** (1993), 81-125.

[9] Kirschner, D. and Webb, G. F., *Immunotherapy of HIV-1 Infection*, Journal of Biological Systems, Vol.6, No.1 (1998) 71-83.

[10] Lukes, D. L., *Differential Equations: Classical to Controlled* Mathematics in Science and Engineering, Academic Press, New York, 1982.

[11] *New Hopes in HIV Disease*, Science, Vol. 274, 20 Dec. 1996 pp. 1988-1991.

[12] Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F., *The Mathematical Theory of Optimal Processes*, Vol 4, Gordon and Breach Science Publishers, 1986.

[13] Tan, Wai-Yuan and Xiang, Z., *Some State Space Model of HIV Pathogenesis Under Treatment by Anti-Viral Drug in HIV- Infected Individuals*, Math. Biosci., 156(1999), No 1-2, 69-94.

Part III

Solving a Parabolic Inverse Problem of Identification Type by an Optimal Control Method

1. Introduction

We identify the "unknown" coefficient of the interaction term of a parabolic system with a Neumann boundary condition in a multidimensional bounded domain.

The solution of the system will represent population concentrations of prey u and predator v. The system has local interaction terms representing a predator-prey situation. The prey equation has Lotka-Voltera (logistic) type growth term. Our aim is to identify the interaction coefficient by optimal control problem techniques.

The state system is the following:

$$\begin{aligned} u_t - b_1\Delta u &= (a_1(x,t) - d(x,t)u)u - h(x)uv \quad \text{in } Q = \Omega \times (0,T) \\ v_t - b_2\Delta v &= -a_2(x,t)v + h(x)uv \end{aligned} \tag{1.1}$$

with initial and boundary conditions:

$$\begin{aligned} u(x,0) &= u_0(x), \quad v(x,0) = v_0(x) && \text{for } x \in \Omega \\ \frac{\partial u}{\partial \nu} &= 0, \quad \frac{\partial v}{\partial \nu} = 0 && \text{on } \partial\Omega \times (0,T). \end{aligned}$$

The coefficients and terms can be interpreted as :

$h(x)$ = coefficient of the interaction term to be identified.

$u(x,t)$ = prey population (first state variable).

$v(x,t)$ = predator population (second state variable).

b_1, b_2 = diffusion coefficients.

a_1, a_2, d = standard logistic growth terms.

$\nu(x)$ = outward unit normal vector at x in $\partial\Omega$.

Our class of admissible interaction coefficients is

$$U = \{h \in L^2(\Omega) : 0 \le h \le M \text{ a. e. in } \Omega\}$$

We seek to minimize the objective functional over $h \in U$:

$$J(h) = \int_W [(u(h) - z_1)^2 + (v(h) - z_2)^2]\, dx\, dt, \tag{1.2}$$

where $W \subset Q$, the measure of W is positive, and z_1, z_2 are observations of the prey and predator populations. We denote the dependence of u and v (solutions of (1.1)) on the interaction coefficient by

$$u = u(h), \quad v = v(h).$$

Given partial (and perhaps noisy) observations z_1, z_2 of the true solution u, v in a subdomain W of Q, we seek to "identify" $h(x)$.

To identify the actual h^*, treat h as a control and choose h such that corresponding u, v are close to given observations z_1, z_2, and an approximate objective functional is minimized. The approximate objective functional for $\beta > 0$ is:

$$J_\beta(h) = \frac{1}{2}\int_W [(u(h) - z_1)^2 + (v(h) - z_2)^2]\, dx\, dt\, + \frac{\beta}{2}\int_\Omega h(x)^2\, dx\,. \tag{1.3}$$

The second term in the right hand side (RHS) is "artificial". This method of solving this identification problem is based on Tikhonov's regularization [**3, 28**]. Tikhonov's approach and its variants seek to minimize the functional $J_\beta(h)$ for a fixed β over a set of unknown data. The function h_β which achieves the minimization of $J_\beta(h)$ for a small β represents the approximate solution of the inverse problem.

For each $\beta > 0$, one considers the unknown data, h, as a control which belongs to U; the control has to be adjusted in order to minimize the functional $J_\beta(h)$ [**18**]. We seek to find the minimum of the approximate objective functional over h in U :

$$J_\beta(h_\beta) = \inf_{h \in U} J_\beta(h).$$

Letting the sequence of β tend toward zero, we will verify that the corresponding sequence, h_β, converges in an appropriate sense to the desired coefficient.

Lenhart, Protopopescu and Yong [**18, 19, 20**] proposed a new approach to the inverse problem of identification, based on optimal control for operator equations as developed by Lions [**23**]. This approach has the following advantages over the classical regularization methods (Tikhonov coupled with optimization schemes [**28**]): (i) a systematic procedure to solve inverse problems of identification type, (ii) an explicit expression for the approximations of the solutions; and (iii) a convenient numerical solution of these approximations. Lenhart et. al. [**1**] applied their approach to solve two examples of identification for hyperbolic inverse problems.

In a parabolic problem, Choulli and Yamamoto recover the coefficient of the zeroth order term when the observation is the final time data by using the strong maximum principle, Hölder a *priori* estimates and the analytic Fredholm alternative [**6**]. In [**12**], Imanuvilov and Yamamoto determine the source term using overdetermined final time data. In [**29**], Yamamoto determined a force term uniquely from the boundary data and estimated the spatial part of force term.

Elayyan proved the uniqueness of a discontinuous diffusion coefficient in a one space dimension inverse parabolic problem using singular solutions of parabolic equations [**7**]. Engl and Zou investigated the stability and convergence rates of the least square method with Tikhonov regularization for the identification of the conductivity distribution in a heat conducting system [**8**]. Isakov identified the coefficients and the source term with the given initial and boundary data and observations at the final time in a parabolic equation [**13**]. Anikonov and Lorenzi reduced the multidimensional nonlinear identification problem related to a general parabolic equation with separated vector variables, containing two unknown coefficients each of which depends on one vector variable only, to the analysis of an infinitely dimensional nonlinear Cauchy problem, containing additional information. Using this procedure they also proved the existence, uniqueness, and stability of the solution of their identification problem [**2**]. In [**5**], Chavent and Kunish formulated the convergence and rate-of-convergence results for non-linear constrained ill-posed problems. Different sufficient conditions for identifiability are studied by Chavent in [**4**]. Both of these papers use least square methods but not the optimal control approach.

Similar techniques of control for parabolic systems are used in various recent publications. See Fister [**10**] and Lenhart et. al. [**16**] for interacting population models. See Leung and Stojanovic [**22**] for an optimal control of elliptic Volterra-Lotka type equations. Leung [**21**] considers optimal control of the harvesting of a prey-predator system in an environment. His species are assumed to be in steady state under diffusion and Volterra-Lotka type of interaction.

Note that the technique given here can be generalized to identify other unknown coefficients of lower terms in this system and in other more general parabolic systems (including variable diffusion and convection). Identification of this interaction coefficient was suggested by Chris Cosner since this coefficient is often difficult to identify from population data.

In section 2, for $\beta > 0$, we discuss the existence and uniqueness of an optimal control, which is an approximation of the unknown coefficient. The optimality system, which characterizes the approximate coefficients, is derived in section 3. Letting $\beta \to 0$, the approximate coefficients converge to a solution of the identification problem in a sense described in section 4. Finally in section 5, we illustrate a numerical example for our problem.

2. Existence of an Optimal Control

The following assumptions are made throughout this paper.

(1) Ω is a smooth bounded domain in $\mathbb{R}^n$.

(2) $u_0(x), v_0(x) \in L^\infty(\Omega)$.

(3) $0 < u_0(x), v_0(x) < B$ for some $B \in \mathbb{R}$.

(4) $a_1, a_2, d \in L^\infty(Q)$, with $a_1 \geq 0, a_2 \geq 0, d \geq 0$, a. e. .

(5) b_1, b_2 are positive constants.

The underlying solution space for system (1.1) is $V = L^2(0, T; H^1(\Omega))$.

DEFINITION 2.1. *We say a pair of functions* $u, v \in L^2(0, T; H^1(\Omega))$, *with* $u_t, v_t \in L^2(0, T; (H^1(\Omega))^*)$, *is a weak solution of (1.1) with given boundary and initial conditions* *provided*

$$\int_0^T < u_t, \phi > \, dt + b_1 \int_Q \nabla u \nabla \phi \, dx \, dt$$
$$= \int_Q a_1 u \phi \, dx \, Lt - \int_Q d u^2 \phi \, dx \, Lt - \int_Q h(x) u v \phi \, dx \, dt \tag{2.1}$$
$$\int_0^T < v_t, \psi > \, dt + b_2 \int_Q \nabla v \nabla \psi \, dx \, dt = - \int_Q a_2 v \psi \, dx \, dt + \int_Q h(x) u v \psi \, dx \, dt$$

for all ϕ, $\psi \in L^2(0, T; H^1(\Omega))$ *and a.e.* $0 \le t \le T$*, where the* $<\ ,\ >$ *inner product is the duality between* $(H^1(\Omega))^*$ *and* $H^1(\Omega)$*, and*

$$u(x,0) = u_0(x), \quad v(x,0) = v_0(x) \tag{2.2}$$

Remark: Since $u, v \in C(0, T; L^2(\Omega))$ from Evans [**9**], the initial conditions (2.2) make sense.

We will prove an existence and uniqueness result for the state system (1.1). This result will be established in Theorem 2.1. First we state and briefly prove a maximum principle for parabolic problems [**25**].

LEMMA 2.1. *Let* $a \in L^\infty(Q)$ *with* $a \ge 0$*. If* u *satisfies* $u \in C^{2,1}(Q) \cap C(\bar{Q})$ *and*

$$\begin{aligned} u_t - b_1 \Delta u &= au && \text{in } Q \\ \frac{\partial u}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ u(x,0) &= u_0(x) && \text{for } x \in \Omega. \end{aligned}$$

Then $\quad 0 \le u \le \left(\sup_\Omega u_0(x) \right) e^{T\|a\|_{L^\infty(\Omega)}}$ *in* $\bar{Q}$*.*

Note: Here $C^{2,1}$ means continuous second derivatives in x and first derivative in t.

Proof: Let $\lambda = \sup_Q a$. By assumption on a, $\lambda \ge 0$. Let $u = e^{\lambda t} w$, then w satisfies

$$\begin{aligned} w_t - b_1 \Delta w + (\lambda - a) w &= 0 && \text{in } Q \\ \frac{\partial w}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ w(x,0) &= u_0(x) && \text{for } x \in \Omega. \end{aligned}$$

By the Parabolic Maximum Principle [**25**]

$$\sup_{\partial\Omega\times(0,T)} w^- \leq w \leq \sup_{\partial\Omega\times(0,T)} w^+$$

where on $\partial\Omega \times (0,T)$, $w^+ = \max(u_0, 0) = u_0, \quad w^- = -\min(u_0, 0) = 0$.

Thus we have $0 \leq w \leq \max_{\partial\Omega\times(0,T)} u_0$, and then we conclude

$$0 \leq u \leq e^{\lambda T} \max_{\Omega} u_0 = \max_{\Omega} u_0(x) e^{T\|a\|_{L^\infty(Q)}} = \|u_0\| e^{T\|a\|_{L^\infty(Q)}}. \quad \square$$

THEOREM 2.1. *Given $h \in U$, there exists a unique solution (u, v) in $V \times V$ solving (1.1).*

Proof: We consider the following problems

$$\begin{aligned} \bar{U}_t - b_1 \Delta \bar{U} &= a_1 \bar{U} && \text{in } Q \\ \frac{\partial \bar{U}}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ \bar{U}(x,0) &= u_0(x) && \text{for } x \in \Omega. \end{aligned} \tag{2.3}$$

$$\begin{aligned} \bar{V}_t - b_2 \Delta \bar{V} &= h(x) \bar{U} \bar{V} && \text{in } Q \\ \frac{\partial \bar{V}}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ \bar{V}(x,0) &= v_0(x) && \text{for } x \in \Omega. \end{aligned} \tag{2.4}$$

The functions $\bar{U}, \bar{V}$ are supersolutions of first and second equations in (1.1) respectively.

To obtain the existence, we will construct two sequences by means of iteration. To obtain L^∞ bounds, we need supersolutions and subsolutions for the u, v iterates.

We first modify the state system in the following form:

$$\begin{aligned} u_t - b_1 \Delta u + Ru &= Ru + (a_1 - du)u - h(x)uv && \text{in } Q \\ \frac{\partial u}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ u(x,0) &= u_0(x) && \text{for } x \in \Omega \\ v_t - b_2 \Delta v + Rv &= Rv - a_2 v + h(x)uv && \text{in } Q \\ \frac{\partial v}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ v(x,0) &= v_0(x) && \text{for } x \in \Omega, \end{aligned} \tag{2.5}$$

where R is a constant satisfying

$$R > \sup_Q \{a_2 + 2d\bar{U} + h\bar{V}\}.$$

By an extension of Lemma 2.1 to weak solutions [**14**], $\bar{U}$ and $\bar{V}$ are L^∞ bounded in Q,

$$\|\bar{U}\| \leq \|u_0\|_{L^\infty(\Omega)} e^{\gamma T} \tag{2.6}$$

$$\|\bar{V}\| \leq \|v_0\|_{L^\infty(\Omega)} e^{\eta T}, \tag{2.7}$$

where $\gamma = \|a_1\|_{L^\infty(\Omega)}$ and $\eta = \|h(x)\|_{L^\infty(Q)} \|u_0\|_{L^\infty(\Omega)} e^{\|a_1\|_{L^\infty(Q)} T}$.

Define $u^1 = \bar{U}$, $v^2 = \bar{V}$, $u^0 = 0$ and $v^1 = 0$, where the superscripts denote the iteration step. For $i = 2, 3, \dots,$ we define u^i and v^{i+1} as the solutions of the following problems, respectively:

$$\begin{aligned} u_t^i - b_1 \Delta u^i + R u^i &= f(u^{i-2}, v^i) && \text{in } Q \\ \frac{\partial u^i}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ u^i(x,0) &= u_0(x) && \text{for } x \in \Omega \end{aligned} \tag{2.8}$$

$$\begin{aligned} v_t^i - b_2 \Delta v^i + R v^i &= g(u^{i-1}, v^{i-2}) && \text{in } Q \\ \frac{\partial v^i}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ v^i(x,0) &= v_0(x) && \text{for } x \in \Omega \end{aligned} \tag{2.9}$$

where $f(u^{i-2}, v^i) = R u^{i-2} + (a_1 - d u^{i-2}) u^{i-2} - h(x) u^{i-2} v^i$

$g(u^{i-1}, v^{i-2}) = R v^{i-2} - a_2 v^{i-2} + h(x) u^{i-1} v^{i-2}.$

Since both the problems (2.8) and (2.9) are linear PDE problems, the solutions $u^i,\ v^i,\ i = 1, 2, \dots,$ exist.

Claim 1

For $i = 1, 2, \dots,$ $\quad 0 \leq u^i \leq \bar{U}, \quad 0 \leq v^i \leq \bar{V}$ in Q.

Proof of Claim 1:

We prove this claim by using the induction method. Since $u^1 = \bar{U}$ and $v^1 = 0$, the case $i = 1$ is trivial. We assume that

$$0 \le u^j \le \bar{U}, \quad 1 \le j \le i-1$$

$$0 \le v^j \le \bar{V}, \quad 1 \le j \le i$$

for some $i > 1$. Then for u^i,

$$\begin{aligned} u_t^i - b_1 \Delta u^i + R u^i &= f(u^{i-2}, v^i) \\ &\ge u^{i-2}\{R - d u^{i-2} - h(x) v^i\} \\ &\ge 0, \end{aligned}$$

Similarly for v^{j+1} we have

$$\begin{aligned} v_t^{j+1} - b_2 \Delta v^{j+1} + R v^{j+1} &= g(u^j, v^{j-1}) \\ &= v^{j-1}\{R - a_2 + h(x) u^j\} \ge v^{j-1}\{R - a_2\} \\ &\ge 0, \end{aligned}$$

since $R > \sup_Q \{a_2 + 2d\bar{U} + h\bar{V}\}$. Then by the parabolic maximum principle

$$u^i \ge \inf_Q u^i \ge -\inf_{\Gamma_T} (u^i)^- = 0.$$

$$v^{j+1} \ge \inf_Q v^{j+1} \ge -\inf_{\Gamma_T} (v^{j+1})^- = 0.$$

where $\Gamma_T = \partial\Omega \times (0, T) \cup \{t = 0\} \times \Omega$, the parabolic boundary of Q.

After subtracting

$$u_t^i - b_1 \Delta u^i + R u^i = f(u^{i-2}, v^i)$$

from (replace $\bar{U}$ by u^1 in (2.3))

$$u_t^1 - b_1 \Delta u^1 + R u^1 = R u^1 + a_1 u^1$$

we get

$$
\begin{aligned}
&(u^1 - u^i)_t - b_1\Delta(u^1 - u^i) + R(u^1 - u^i) \\
&= R(u^1 - u^{i-2}) + a_1(u^1 - u^{i-2}) + d(u^{i-2})^2 + h(x)u^{i-2}v^i.
\end{aligned}
$$

As $\bar{U}$ is a supersolution, hence $u^{i-2} \leq \bar{U}$ and similarly $v^i \leq \bar{V}$. Since R, a_1, d, h are all positive quantities, thus the RHS of above equation is positive. By applying maximum principle, $0 \leq u^i \leq \bar{U}$. A similar result holds for v^i. Hence by an induction argument the claim 1 holds.

Claim 2

For $u^i \geq 0, v^i \geq 0$, $i = 1, 2, \ldots,$

f is increasing in u^{i-2}, f is decreasing in v^i.

g is increasing in u^{i-1}, g is increasing in v^{i-2}.

Proof of Claim 2:

We calculate the derivatives of f and g to verify the above properties:

$$
\begin{aligned}
f(u^{i-2}, v^i) &= Ru^{i-2} + (a_1 - du^{i-2})u^{i-2} - h(x)u^{i-2}v^i \\
g(u^{i-1}, v^{i-2}) &= Rv^{i-2} - a_2v^{i-2} + h(x)u^{i-1}v^{i-2} \\
\frac{\partial f}{\partial u^{i-2}} &= R + a_1 - 2du^{i-2} - h(x)v^i \geq 0 \quad \text{(by choice of } R\text{)} \\
\frac{\partial f}{\partial v^i} &= -h(x)u^{i-2} < 0 \quad \text{(since both } h \text{ and } u \text{ are } \geq 0\text{)} \\
\frac{\partial g}{\partial u^{i-1}} &= h(x)v^{i-2} > 0 \\
\frac{\partial g}{\partial v^{i-2}} &= R - a_2 + hu^{i-1} > 0 \quad \text{(by choice of } R\text{)}.
\end{aligned}
$$

Hence the claim is established.

Claim 3

There exist $\bar{u}, \bar{v}, u, v$ in V, such that the following monotone pointwise convergence holds:

$$u^{2i} \nearrow \bar{u}, \qquad u^{2i+1} \searrow u \text{ in } Q$$

$$v^{2i} \searrow \bar{v}, \qquad v^{2i+1} \searrow v \text{ in } Q,$$

pointwise for $i = 1, 2, \ldots$.

Proof of Claim 3:

To prove above convergence we use the induction method. Since $v^1 = 0$, we have

$$(v^5 - v^3)_t - b_2\Delta(v^5 - v^3) + R(v^5 - v^3) = v^3(R - a_2 + h(x)u^4) \geq 0,$$

and thus $0 = v^1 \leq v^3 \leq v^5$. Similarly, we have $u^3 \leq u^1, v^2 \geq v^4$, and $u^0 \leq u^2$. Also $u^2 \leq u^3$ and $v^4 \geq v^3$ since

$$\begin{aligned}
(u^2 - u^3)_t - b_1\Delta(u^2 - u^3) + R(u^2 - u^3) &= f(u^0, v^2) - f(u^1, v^3) \\
&= -f(u^1, v^3) \\
&= -[Ru^1 + (a_1 - du^1)u^1 - h(x)u^1v^3] \\
&= -[R + (a_1 - du^1) - h(x)v^3]u^1 \leq 0, \\
(v^4 - v^3)_t - b_2\Delta(v^4 - v^3) + R(v^4 - v^3) &= g(u^3, v^2) - g(u^2, v^1) \\
&= (R - a_2)(v^2 - v^1) + h(x)(u^3v^2 - u^2v^1) \\
&= Rv^2 - a_2v^2 + h(x)u^2v^2 \\
&= [R - a_2 + h(x)u^2]v^2 \geq 0.
\end{aligned}$$

Fix i, assume that for all j, $j \leq i - 1$,

$$u^{2j} \nearrow \qquad u^{2j+1} \searrow$$

$$v^{2j} \searrow \qquad v^{2j+1} \nearrow.$$

For i, we compare v^{2i+2} and v^{2i}. From (2.9), we have

$$
\begin{aligned}
&(v^{2i+2}-v^{2i})_t - b_2\Delta(v^{2i+2}-v^{2i}) + R(v^{2i+2}-v^{2i})\\
&= g(u^{2i+1},v^{2i}) - g(u^{2i-1},v^{2i-2}) \leq 0
\end{aligned}
$$

Then $v^{2i+2} \leq v^{2i}$. Other cases can be proved using similar arguments. Hence by boundedness of u^i, v^i and monotone properties (as proved in claim 2), we get the pointwise convergence.

Claim 4

The subsequences of $\{u^i\}, \{v^i\}$ satisfy

$$u^{2j} \rightharpoonup \bar{u}, \quad u^{2j+1} \rightharpoonup u, \quad v^{2j} \rightharpoonup \bar{v}, \quad v^{2j+1} \rightharpoonup v \quad \text{weakly in } V.$$

Proof of Claim 4:

Consider the RHS of (2.8) and (2.9)

$$
\begin{aligned}
\|f(u^{i-2},v^i)\|_{L^\infty(Q)} &= \|Ru^{i-2} + a_1u^{i-2} - d(u^{i-2})^2 - hu^{i-2}v^i\|_{L^\infty(Q)}\\
&\leq \|u^1\|_{L^\infty(Q)}\big(R + \|a_1\|_{L^\infty(Q)} + \|d\|_{L^\infty(Q)}\|u^1\|_{L^\infty(Q)}\\
&\quad +\|h(x)\|_{L^\infty(Q)}\|v^2\|_{L^\infty(Q)}\big)\\
&\equiv C_1\\
\|g(u^{i-1},v^{i-2})\|_{L^\infty(Q)} &= \|Rv^{i-2} + a_2v^{i-2} + hu^{i-1}v^{i-2}\|_{L^\infty(Q)}\\
&\leq \|v^2\|_{L^\infty(Q)}\big(R + \|a_2\|_{L^\infty(Q)} + \|h(x)\|_{L^\infty(Q)}\|u^1\|_{L^\infty(Q)}\big)\\
&\equiv C_2.
\end{aligned}
$$

Thus RHS of (2.8) and (2.9) are bounded in $L^\infty(Q)$, then using u^k as test function in its own equation in (2.8) and using *a priori* L^∞ bounds on u^k, v^k, we obtain

$$
\begin{aligned}
&\sup_{0\leq t\leq T}\left(\int_\Omega (u^k)^2(x,t)\,dx\right) + \int_0^T\int_\Omega b_1|\nabla u^k|^2\,dx\,dt\\
&\leq C_3\int_\Omega u_0^2\,dx + C_4\int_0^T\int_\Omega \left|f(u^{k-2},v^k)\right|\,dx\,dt
\end{aligned}
$$

$$\le C_3 \int_\Omega u_0^2\,dx + C_4 C_1 T\, meas(\Omega).$$

Thus the sequences $\{u^{2i}\}, \{u^{2i+1}\}$ are uniformly bounded in V. Similarly v^{2i}, v^{2i+1} are also bounded in V.

Hence using the weak compactness of the sequences in V, we get

$$u^{2j} \rightharpoonup \bar{u}, \quad u^{2j+1} \rightharpoonup u, \quad v^{2j} \rightharpoonup \bar{v}, \quad v^{2j+1} \rightharpoonup v \quad \text{weakly in } V.$$

Since we have pointwise convergence on each sequence by claim 3, this weak convergence is also on the whole (even or odd) sequence (not just on a subsequence).

Claim 5

The subsequences of $\{u^i\}, \{v^i\}$ satisfy

$$u^{2j} \to \bar{u}, \quad u^{2j+1} \to u, \quad v^{2j} \to \bar{v}, \quad v^{2j+1} \to v \quad \text{strongly in } L^2(Q).$$

Proof of Claim 5:

Using the V boundedness on $u^{2i}, u^{2i+1}, v^{2i}, v^{2i+1}$ and the PDE in (2.8) and (2.9), we obtain that

$$\|u_t^{2i}\|, \quad \|u_t^{2i+1}\|, \quad \|v_t^{2i}\|, \quad \|v_t^{2i+1}\| \text{ are bounded in } L^2(0,T;(H^1(\Omega))^*).$$

Hence, using weak compactness again, we have

$$u_t^{2i} \rightharpoonup \bar{u}_t, \quad u_t^{2i+1} \rightharpoonup u_t, \quad v_t^{2i} \rightharpoonup \bar{v}_t, \quad v_t^{2i+1} \rightharpoonup v_t \text{ weakly in } L^2(0,T;(H^1(\Omega))^*).$$

Using a compactness result by Simon [**26**], we have

$$u^{2i} \to \bar{u}, \quad u^{2i+1} \to u, \quad v^{2i} \to \bar{v}, \quad v^{2i+1} \to v \quad \text{strongly in } L^2(Q).$$

Passing to the limit in the u^{2i}, u^{2i+1}, v^{2i} and v^{2i+1} PDE's, we obtain:

$$\begin{aligned} \bar{u}_t - b_1 \Delta \bar{u} &= a_1 \bar{u} - d\bar{u}^2 - h(x)\bar{u}\bar{v} && \text{in } Q \\ \frac{\partial \bar{u}}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ \bar{u}(x,0) &= u_0(x) && \text{for } x \in \Omega \end{aligned} \tag{2.10}$$

$$\begin{aligned} u_t - b_1\Delta u &= a_1 u - du^2 - h(x)uv && \text{in } Q \\ \frac{\partial u}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ u(x,0) &= u_0(x) && \text{for } x \in \Omega \end{aligned} \tag{2.11}$$

$$\begin{aligned} \bar v_t - b_2\Delta \bar v &= -a_2 \bar v + h(x)u\bar v. && \text{in } Q \\ \frac{\partial \bar v}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ \bar v(x,0) &= v_0(x) && \text{for } x \in \Omega \end{aligned} \tag{2.12}$$

$$\begin{aligned} v_t - b_2\Delta v &= -a_2 v + h(x)\bar u v. && \text{in } Q \\ \frac{\partial v}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ v(x,0) &= v_0(x) && \text{for } x \in \Omega. \end{aligned} \tag{2.13}$$

In order to prove uniqueness, we need to show $\bar u = u,\ v = \bar v$. Let $u = e^{\lambda t} w$, $\bar u = e^{\lambda t}\bar w$, $v = e^{\lambda t} z$ and $\bar v = e^{\lambda t}\bar z$, where $\lambda > 0$ is to be chosen.

We give as an example the transformed PDE problem (2.11) for the $u = e^{\lambda t} w$ case.

$$\begin{aligned} w_t - b_1\Delta w + \lambda w &= a_1 w - de^{\lambda t} w^2 - h(x)e^{\lambda t} wz && \text{in } Q \\ \frac{\partial w}{\partial \nu} &= 0 && \text{on } \partial\Omega \times (0,T) \\ w(x,0) &= u_0(x) && \text{for } x \in \Omega. \end{aligned} \tag{2.14}$$

We consider the weak formulation of the $w - \bar w$ and $z - \bar z$ problem. Using $w - \bar w$ and $z - \bar z$ as test functions and after adding both weak formulations, we obtain:

$$\begin{aligned} &\int_Q \{(w-\bar w)_t(w-\bar w) + b_1(|\nabla(w-\bar w)|)^2 + \lambda(w-\bar w)^2\}\,dx\,dt \\ &\quad + \int_Q \{(z-\bar z)_t(z-\bar z) + b_2(|\nabla(z-\bar z)|)^2 + \lambda(z-\bar z)^2\}\,dx\,dt \\ &= \int_Q \{a_1(w-\bar w)^2 - de^{\lambda t}(w^2-\bar w^2)(w-\bar w) - h(x)e^{\lambda t}(w-\bar w)(wz-\bar w\,\bar z)\}\,dx\,dt \\ &\quad + \int_Q \{-a_2(z-\bar z)^2 + h(x)e^{\lambda t}(\bar w z - w\bar z)(z-\bar z)\}\,dx\,dt. \end{aligned}$$

Using $wz - \bar w\,\bar z = z(w-\bar w) + \bar w(z-\bar z)$, $\bar w z - w\bar z = \bar w(z-\bar z) - \bar z(w-\bar w)$, $w(x,0) = u_0(x) = \bar w(x,0)$, and

$$\int_Q (w-\bar w)_t(w-\bar w)\,dx\,dt \;=\; \frac{1}{2}\int_Q \frac{d}{dt}(w-\bar w)^2\,dx\,dt$$

$$= \quad \frac{1}{2}\int_\Omega (w(x,T)-\bar{w}(x,T))^2\,dx,$$

we obtain

$$\begin{aligned}
&\frac{1}{2}\int_\Omega \{[w(x,T)-\bar{w}(x,T)]^2+[z(x,T)-\bar{z}(x,T)]^2\}\,dx \\
&\quad+\int_Q [b_1(|\nabla(w-\bar{w})|)^2+b_2(|\nabla(z-\bar{z})|)^2]\,dx\,dt+\lambda\int_Q[(w-\bar{w})^2+(z-\bar{z})^2]\,dx\,dt \\
&=\int_Q [a_1(w-\bar{w})^2-a_2(z-\bar{z})^2]\,dx\,dt+\int_Q e^{\lambda t}(w-\bar{w})^2[-d(w+\bar{w})-h(x)z]\,dx\,dt \\
&\quad+\int_Q e^{\lambda t}(z-\bar{z})^2h(x)\bar{w}\,dx\,dt+\int_Q h(x)e^{\lambda t}(w-\bar{w})(z-\bar{z})(-\bar{w}-\bar{z})\,dx\,dt.
\end{aligned}$$

Using the boundedness of $w,\bar{w},z,\bar{z}$ (via $u,\bar{u},v,\bar{v}$), a_1,a_2,d,h and Cauchy inequality, we obtain:

$$\begin{aligned}
&\frac{1}{2}\int_\Omega \{[w(x,T)-\bar{w}(x,T)]^2+[z(x,T)-\bar{z}(x,T)]^2\}\,dx \\
&\quad+\int_Q \{[b_1(|\nabla(w-\bar{w})|)^2+b_2(|\nabla(z-\bar{z})|)^2]\,dx\,dt \\
&\quad+(\lambda-C)\int_Q[(w-\bar{w})^2+(z-\bar{z})^2]\,dx\,dt \\
&\le 0,
\end{aligned}\tag{2.15}$$

where C depends on the coefficients, T and the solution bounds.

If we choose $\lambda > C$, then inequality (2.15) holds if and only if

$$w=\bar{w}, \qquad z=\bar{z} \quad \text{a.e. in } Q.$$

Therefore, $u=\bar{u}$ and $v=\bar{v}$ a.e. in Q, and u,v solve the state system (1.1). Hence the solution to the state system (1.1) exists and the uniqueness of u,v follows similarly as in the above argument. $\square$

Now we present the main result of this section.

THEOREM 2.2. *There exists an optimal control in U that minimizes the functional $J_\beta(h)$.*

Proof: By the form of $J_\beta(h)$, $\inf\{J_\beta(h)\,|\,h \in U\} \geq 0$. Then there exists a minimizing sequence $h^n \in U$ such that

$$\lim_{n\to\infty} J_\beta(h^n) = \inf\{J_\beta(h)\,|\,h \in U\}.$$

By the existence and uniqueness of solutions to the state system (1.1), we define

$$u^n = u(h^n) \qquad v^n = v(h^n) \quad \text{for each } n.$$

We consider for $0 \leq t \leq T$

$$\begin{aligned}&\int_{\Omega\times(0,t)} \left[u_t^n u^n + b_1|\nabla u^n|^2 - a_1(u^n)^2 + v_t^n v^n + b_2|\nabla v^n|^2 + a_2(v^n)^2\right]\,dx\,d\tau \\ &= -\int_{\Omega\times(0,t)} \left[h(x)(u^n)^2 v^n - d(u^n)^2 u^n + h(x)u^n(v^n)^2\right]\,dx\,d\tau.\end{aligned} \tag{2.16}$$

Using Cauchy inequality and the fact that state variables are uniformly bounded, we obtain:

$$\begin{aligned}&\frac{1}{2}\int_\Omega \left([u^n(x,t)]^2 + [v^n(x,t)]^2\right)\,dx + \int_{\Omega\times(0,t)} \left[b_1|\nabla u^n|^2 + b_2|\nabla v^n|^2\right]\,dx\,d\tau \\ &\leq \frac{1}{2}\int_\Omega [(u_0)^2 + (v_0)^2]\,dx + C\int_{\Omega\times(0,t)} [(u^n)^2 + (v^n)^2]\,dx\,d\tau.\end{aligned} \tag{2.17}$$

In particular,

$$\begin{aligned}\int_\Omega \left([u^n(x,t)]^2 + [v^n(x,t)]^2\right)\,dx \leq &\int_\Omega [(u_0)^2 + (v_0)^2]\,dx \\ &+ 2C\int_{\Omega\times(0,t)} [(u^n)^2 + (v^n)^2]\,dx\,d\tau.\end{aligned} \tag{2.18}$$

After applying the Gronwall's inequality to (2.18), we get :

$$\begin{aligned}\int_\Omega [(u^n)^2 + (v^n)^2](x,t)\,dx &\leq \left(1 + 2\,C\,t\,e^{2Ct}\right)\int_\Omega (u_0^2 + v_0^2)\,dx \\ &\leq \left(1 + 2\,C\,T\,e^{2\,C\,T}\right)\int_\Omega (u_0^2 + v_0^2)\,dx.\end{aligned} \tag{2.19}$$

Using that estimate with (2.17) gives

$$\sup_{0\leq t\leq T}\left\{\int_\Omega \left([u^n(x,t)]^2 + [v^n(x,t)]^2\right)\,dx\right\} + \int_{\Omega\times(0,T)} \left[b_1|\nabla u^n|^2 + b_2|\nabla v^n|^2\right]\,dx\,d\tau$$

$$\le \int_\Omega [(u_0)^2 + (v_0)^2]\, dx + 2\,C \int_{\Omega\times(0,T)} [(u^n)^2 + (v^n)^2]\, dx\, d\tau$$
$$\le \int_\Omega [(u_0)^2 + (v_0)^2]\, dx + 2\,C\,T \left[1 + 2C\,T\, e^{2CT}\right] \int_\Omega [(u_0)^2 + (v_0)^2]\, dx$$
$$= \left[1 + 2\,C\,T \left(1 + 2\,C\,T e^{2CT}\right)\right] \int_\Omega [(u_0)^2 + (v_0)^2]\, dx. \tag{2.20}$$

On a subsequence, as $n \to \infty$, $h^n \to h_\beta$ in $L^2(Q)$. As in the proof of the existence of solutions to the state system (1.1) (mainly claim 5), u_t^n and v_t^n are bounded and weakly convergent in $L^2(0,T;(H^1(\Omega))^*)$, and

$$u^n \to u_\beta, \quad v^n \to v_\beta \ \text{ strongly in } L^2(Q) \text{ and weakly in } V. \tag{2.21}$$

We further note that:

$$\int_Q |(h^n u^n v^n - h_\beta u_\beta v_\beta)\,\phi| \ dx\, dt$$
$$\le \int_Q |h^n u^n v^n - h^n u^n v_\beta + h^n u^n v_\beta - h^n u_\beta v_\beta + h^n u_\beta v_\beta - h_\beta u_\beta v_\beta| \ |\phi|\, dx\, dt$$
$$\to 0 \quad \text{as } n \to \infty,$$

since

$$\int_Q |u^n - u_\beta||h^n v_\beta||\phi|\, dx\, dt \le \left(\int_Q (u^n - u_\beta)^2\, dx\, dt\right)^{\frac{1}{2}} C_{h,v} \left(\int_Q \phi^2\, dx\, dt\right)^{\frac{1}{2}}$$
$$\to 0 \text{ as } n \to \infty,$$
$$\int_Q |v^n - v_\beta||h^n u^n||\phi|\, dx\, dt \le \left(\int_Q (v^n - v_\beta)^2\, dx\, dt\right)^{\frac{1}{2}} C_{h,u} \left(\int_Q \phi^2\, dx\, dt\right)^{\frac{1}{2}}$$
$$\to 0 \text{ as } n \to \infty, \quad \text{and}$$
$$\left|\int_Q (h^n - h_\beta) u_\beta v_\beta \phi\, dx\, dt\right| \to 0 \text{ as } n \to \infty, \quad \text{by weak convergence of } h^n \text{ sequence.}$$

Thus passing to the limit in the u^n, v^n PDE's and using the convergences as in Theorem 2.1, we have that (u_β, v_β) is weak solution of (1.1) associated with h_β. Thus we have

$$J_\beta(h_\beta) \ \ge \inf\{J_\beta(h^n) \,|\, h^n \in U\}. \tag{2.22}$$

For the weakly convergent control sequence [**9**],

$$\int_Q (h_\beta)^2\,dx\,dt \le \underline{\lim}_{n\to\infty} \int_Q (h^n)^2\,dx\,dt. \tag{2.23}$$

Hence (2.21) and (2.23) imply

$$J_\beta(h_\beta) = \min_{h\in U} J_\beta(h). \tag{2.24}$$

Therefore h_β is an optimal control that minimizes the objective functional. □

3. Derivation of the Optimality System

We now derive the optimality system which consists of the state system coupled with the adjoint system. In order to obtain the necessary conditions for the optimality system we differentiate the objective functional with respect to the control. As our objective functional also depends on u and v, we differentiate u and v with respect to the control h.

THEOREM 3.1. *The mapping $h \in U \to (u,v) \in V \times V$ is differentiable in the following sense:*

$$\begin{aligned} \frac{u(h+\varepsilon k)-u(h)}{\varepsilon} &\rightharpoonup \psi_1 \\ \frac{v(h+\varepsilon k)-v(h)}{\varepsilon} &\rightharpoonup \psi_2 \end{aligned}$$

weakly in V as $\varepsilon \to 0$ for any $h \in U$ and $k \in L^\infty(Q)$ s.t. $(h+\varepsilon k) \in U$ for ε small. Also ψ_1, ψ_2 satisfy the following system:

$$\begin{aligned} (\psi_1)_t - b_1\Delta\psi_1 &= a_1\psi_1 - 2d\psi_1 u - h(u\psi_2 + v\psi_1) - kuv && \text{in } Q \\ (\psi_2)_t - b_2\Delta\psi_2 &= -a_2\psi_2 + h(u\psi_2 + v\psi_1) + kuv && \\ \psi_1(x,0) &= 0 \quad = \psi_2(x,0) && \text{for } x \in \Omega \\ \frac{\partial\psi_1}{\partial\nu} &= 0 \quad = \frac{\partial\psi_2}{\partial\nu} && \text{on } \partial\Omega \times (0,T) \end{aligned} \tag{3.1}$$

Proof: Define $u^\varepsilon = u(h+\varepsilon k)$, $v^\varepsilon = v(h+\varepsilon k)$, $u = u(h)$, and $v = v(h)$. We do a change of variables: $u^\varepsilon = e^{\lambda t}w^\varepsilon$, $u = e^{\lambda t}w$, $v^\varepsilon = e^{\lambda t}z^\varepsilon$, $v = e^{\lambda t}z$, where $\lambda > 0$ is to be chosen below. We have

$$u_t^\varepsilon = e^{\lambda t}\left(w_t^\varepsilon + \lambda w^\varepsilon\right) \qquad v_t^\varepsilon = e^{\lambda t}\left(z_t^\varepsilon + \lambda z^\varepsilon\right)$$
$$u_t = e^{\lambda t}\left(w_t + \lambda w\right) \qquad v_t = e^{\lambda t}\left(z_t + \lambda z\right).$$

Using (1.1), we obtain

$$(u^\varepsilon - u)_t - b_1\Delta\left(u^\varepsilon - u\right) = a_1\left(u^\varepsilon - u\right) - d\left(\left(u^\varepsilon\right)^2 - u^2\right) - (h+\varepsilon k)u^\varepsilon v^\varepsilon + huv$$
$$(v^\varepsilon - v)_t - b_2\Delta\left(v^\varepsilon - v\right) = -a_2\left(v^\varepsilon - v\right) + (h+\varepsilon k)u^\varepsilon v^\varepsilon - huv.$$

In terms of w^ε, w variables, we have

$$e^{\lambda t}\left[\frac{(w^\varepsilon - w)_t}{\varepsilon} + \lambda\frac{(w^\varepsilon - w)}{\varepsilon} - b_1\Delta\frac{(w^\varepsilon - w)}{\varepsilon}\right]$$
$$= e^{\lambda t}\left[\frac{(w^\varepsilon - w)}{\varepsilon}a_1 - d\frac{e^{\lambda t}\left(\left(w^\varepsilon\right)^2 - w^2\right)}{\varepsilon} - he^{\lambda t}\left(\frac{w^\varepsilon z^\varepsilon - wz}{\varepsilon}\right) - ke^{\lambda t}w^\varepsilon z^\varepsilon\right]$$

and

$$e^{\lambda t}\left[\frac{(z^\varepsilon - z)_t}{\varepsilon} + \lambda\frac{(z^\varepsilon - z)}{\varepsilon} - b_2\Delta\frac{(z^\varepsilon - z)}{\varepsilon}\right]$$
$$= e^{\lambda t}\left[-a_2\left(\frac{z^\varepsilon - z}{\varepsilon}\right) + he^{\lambda t}\left(\frac{w^\varepsilon z^\varepsilon - wz}{\varepsilon}\right) + ke^{\lambda t}w^\varepsilon z^\varepsilon\right].$$

Using the weak formulation with the test functions $\dfrac{(w^\varepsilon - w)}{\varepsilon}$ and $\dfrac{(z^\varepsilon - z)}{\varepsilon}$, on the set $Q_1 = \Omega \times (0, t_1)$ for $0 < t_1 \le T$, we have

$$\int_{Q_1}\left[\left(\frac{w^\varepsilon - w}{\varepsilon}\right)\left(\frac{w^\varepsilon - w}{\varepsilon}\right)_t + \lambda\left(\frac{w^\varepsilon - w}{\varepsilon}\right)\left(\frac{w^\varepsilon - w}{\varepsilon}\right) + b_1\left|\nabla\left(\frac{w^\varepsilon - w}{\varepsilon}\right)\right|^2\right]dx\,dt$$
$$= \int_{Q_1}\left[a_1\left(\frac{w^\varepsilon - w}{\varepsilon}\right)^2 - d\left(\frac{w^\varepsilon - w}{\varepsilon}\right)^2 e^{\lambda t}\left(w^\varepsilon + w\right)\right.$$
$$\left.-he^{\lambda t}\left(w^\varepsilon\left(\frac{w^\varepsilon - w}{\varepsilon}\right)\left(\frac{z^\varepsilon - z}{\varepsilon}\right) + z\left(\frac{w^\varepsilon - w}{\varepsilon}\right)^2\right) - ke^{\lambda t}w^\varepsilon z^\varepsilon\left(\frac{w^\varepsilon - w}{\varepsilon}\right)\right]dx\,dt$$

and

$$\int_{Q_1}\left[\left(\frac{z^\varepsilon - z}{\varepsilon}\right)\left(\frac{z^\varepsilon - z}{\varepsilon}\right)_t + \lambda\left(\frac{z^\varepsilon - z}{\varepsilon}\right)\left(\frac{z^\varepsilon - z}{\varepsilon}\right) + b_2\left|\nabla\left(\frac{z^\varepsilon - z}{\varepsilon}\right)\right|^2\right] dx\,dt$$
$$= \int_{Q_1}\left[-a_2\left(\frac{z^\varepsilon - z}{\varepsilon}\right)^2 + he^{\lambda t}\left[w^\varepsilon\left(\frac{z^\varepsilon - z}{\varepsilon}\right)^2 + z\left(\frac{w^\varepsilon - w}{\varepsilon}\right)\left(\frac{z^\varepsilon - z}{\varepsilon}\right)\right]\right.$$
$$\left. + ke^{\lambda t}w^\varepsilon z^\varepsilon\left(\frac{z^\varepsilon - z}{\varepsilon}\right)\right] dx\,dt.$$

Adding the equations and estimating the RHS, we obtain

$$\frac{1}{2}\int_{\Omega\times\{t_1\}}\left[\left(\frac{w^\varepsilon - w}{\varepsilon}\right)^2 + \left(\frac{z^\varepsilon - z}{\varepsilon}\right)^2\right] dx + \lambda\int_{Q_1}\left[\left(\frac{w^\varepsilon - w}{\varepsilon}\right)^2 + \left(\frac{z^\varepsilon - z}{\varepsilon}\right)^2\right] dx\,dt$$
$$+ \int_{Q_1}\left[b_1\left|\nabla\left(\frac{w^\varepsilon - w}{\varepsilon}\right)\right|^2 + b_2\left|\nabla\left(\frac{z^\varepsilon - z}{\varepsilon}\right)\right|^2\right] dx\ dt$$
$$= \int_{Q_1}\left(a_1 - de^{\lambda t}(w^\varepsilon + w) - he^{\lambda t}z\right)\left(\frac{w^\varepsilon - w}{\varepsilon}\right)^2 dx\,dt$$
$$- \int_{Q_1}\left(he^{\lambda t}w^\varepsilon - he^{\lambda t}z\right)\left(\frac{w^\varepsilon - w}{\varepsilon}\right)\left(\frac{z^\varepsilon - z}{\varepsilon}\right) dx\,dt$$
$$+ \int_{Q_1}\left(-a_2 + he^{\lambda t}w^\varepsilon\right)\left(\frac{z^\varepsilon - z}{\varepsilon}\right)^2 dx\,dt$$
$$- \int_{Q_1} ke^{\lambda t}w^\varepsilon z^\varepsilon\left[\left(\frac{w^\varepsilon - w}{\varepsilon}\right) - \left(\frac{z^\varepsilon - z}{\varepsilon}\right)\right] dx\,dt$$
$$\leq \left(C_1 + C_2e^{2\lambda t_1}\right)\int_{Q_1}\left[\left(\frac{w^\varepsilon - w}{\varepsilon}\right)^2 + \left(\frac{z^\varepsilon - z}{\varepsilon}\right)^2\right] dx\,dt$$
$$+ \int_{Q_1} k^2\left(w^\varepsilon\right)^2\left(z^\varepsilon\right)^2\ dx\,dt \tag{3.2}$$

Continuing to estimate using L^∞ bounds on the coefficients and $w, z, w^\varepsilon, z^\varepsilon$, we have

$$\frac{1}{2}\int_{\Omega\times t_1}\left[\left(\frac{w^\varepsilon(x,t_1) - w(x,t_1)}{\varepsilon}\right)^2 + \left(\frac{z^\varepsilon(x,t_1) - z(x,t_1)}{\varepsilon}\right)^2\right] dx$$

$$+\int_{Q_1}\left[b_1\left|\nabla\left(\frac{w^\varepsilon-w}{\varepsilon}\right)\right|^2+b_2\left|\nabla\left(\frac{z^\varepsilon-z}{\varepsilon}\right)\right|^2\right]dx\,dt$$

$$+\left(\lambda-\left(C_1+C_2e^{2\lambda t_1}\right)\right)\int_{Q_1}\left[\left(\frac{w^\varepsilon-w}{\varepsilon}\right)^2+\left(\frac{z^\varepsilon-z}{\varepsilon}\right)^2\right]dx\,dt\le C_3\int_{Q_1}k^2\,dx\,dt.$$

For $\lambda > C_1+C_2$ and t_1 small $\left(t_1<\frac{1}{2\lambda}\ln\frac{\lambda-C_1}{C_2}\right)$, we conclude

$$\left\|\frac{w^\varepsilon-w}{\varepsilon}\right\|^2_{L^2(0,t_1;H^1(\Omega))}+\left\|\frac{z^\varepsilon-z}{\varepsilon}\right\|^2_{L^2(0,t_1;H^1(\Omega))}\le C_4\int_Q k^2\,dx\,dt.$$

Similarly this estimate can be carried out on intervals, $[t_1,2t_1],[2t_1,3t_1],\ldots,$ and the estimate finally holds on $[0,T]$. This estimate justifies the convergence of w,z quotients, and hence

$$\frac{u^\varepsilon-u}{\varepsilon}\rightharpoonup\psi_1,\quad\frac{v^\varepsilon-v}{\varepsilon}\rightharpoonup\psi_2\quad\text{weakly in } V \tag{3.3}$$

Similar to Theorem 2.1, we obtain

$$\left(\frac{u^\varepsilon-u}{\varepsilon}\right)_t\rightharpoonup(\psi_1)_t\quad\text{and}\quad\left(\frac{v^\varepsilon-v}{\varepsilon}\right)_t\rightharpoonup(\psi_2)_t\quad\text{in } L^2(0,T;H^1(\Omega)^*),$$

and $\frac{u^\varepsilon-u}{\varepsilon}\to\psi_1\quad\frac{v^\varepsilon-v}{\varepsilon}\to\psi_2$ strongly in $L^2(Q)$. These convergences also give $u^\varepsilon\to u$ and $v^\varepsilon\to v$ strongly in $L^2(Q)$.

To see the system satisfied by ψ_1,ψ_2, consider terms from the system satisfied by $\frac{u^\varepsilon-u}{\varepsilon}$ and $\frac{v^\varepsilon-v}{\varepsilon}$; for example,

$$\begin{aligned}&\frac{1}{\varepsilon}d\left((u^\varepsilon)^2-u^2\right)=d\frac{1}{\varepsilon}\left(u^\varepsilon-u\right)\left(u^\varepsilon+u\right)\\&\to 2d\psi_1u\qquad\text{as}\quad\varepsilon\to0,\\&\frac{1}{\varepsilon}\left[(h+\varepsilon k)u^\varepsilon v^\varepsilon-huv\right]=\frac{1}{\varepsilon}\left[h(u^\varepsilon v^\varepsilon-uv)+\varepsilon k u^\varepsilon v^\varepsilon\right]\\&\to h(u\psi_2+v\psi_1)+kuv\qquad\text{(using (3.3))}\end{aligned}$$

since $u^\varepsilon\to u,\ v^\varepsilon\to v$ as $\varepsilon\to0$.

The above estimates justify passing to the limit in the PDE's, which are satisfied by $\frac{u^\varepsilon-u}{\varepsilon}$ and $\frac{v^\varepsilon-v}{\varepsilon}$, and we conclude that ψ_1,ψ_2 solve (3.1). $\square$

To derive the optimality system and to characterize the pair of optimal controls, we need adjoint variables and adjoints of the operators associated with the ψ_1, ψ_2 system. We write the ψ_1, ψ_2 PDE system as

$$\mathcal{L}\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} -kuv \\ kuv \end{pmatrix}$$

where $$\mathcal{L}\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} = \begin{pmatrix} \mathcal{L}_1\psi_1 \\ \mathcal{L}_2\psi_2 \end{pmatrix} + M\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

$$\begin{pmatrix} \mathcal{L}_1\psi_1 \\ \mathcal{L}_2\psi_2 \end{pmatrix} = \begin{pmatrix} (\psi_1)_t - b_1\Delta\psi_1 \\ (\psi_2)_t - b_2\Delta\psi_2 \end{pmatrix} \quad , \quad M = \begin{pmatrix} -a_1 + 2ud + hv & hu \\ -hv & a_2 - hu \end{pmatrix}.$$

We define the adjoint PDE system as $\mathcal{L}^*\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} u - z_1 \\ v - z_2 \end{pmatrix}\chi_W$

where χ_W is the characteristic function of the set W and

$$\mathcal{L}^*\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \mathcal{L}_1^* p \\ \mathcal{L}_2^* q \end{pmatrix} + M^T\begin{pmatrix} p \\ q \end{pmatrix}$$

with

$$\begin{pmatrix} \mathcal{L}_1^* p \\ \mathcal{L}_2^* q \end{pmatrix} = \begin{pmatrix} -p_t - b_1\Delta p \\ -q_t - b_2\Delta q \end{pmatrix} \tag{3.4}$$

and

$$M^T = \begin{pmatrix} -a_1 + 2ud + hv & -hv \\ hu & a_2 - hu \end{pmatrix}.$$

Thus the adjoint PDE system is

$$\begin{pmatrix} -p_t - b_1\Delta p \\ -q_t - b_2\Delta q \end{pmatrix} + \begin{pmatrix} -a_1 + 2ud + hv & -hv \\ hu & a_2 - hu \end{pmatrix}\begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} u - z_1 \\ v - z_2 \end{pmatrix}\chi_W.$$

The components of the RHS of this system are the derivative of the integrand of the objective functional with respect to each state respectively. Before computing this RHS, the functional (1.3) is rewritten as

$$\frac{1}{2}\int_Q \left[(u(h)-z_1)^2+(v(h)-z_2)^2\right]\chi_W\,dx\,dt\;+\frac{\beta}{2}\int_\Omega h(x)^2\,dx\,.$$

We illustrate in a formal way the appropriate weak sense for solution of the adjoint system:

$$\begin{aligned}
&\int_Q (\psi_1,\psi_2)\,\mathcal{L}^*\begin{pmatrix} p \\ q \end{pmatrix}\,dx\,dt = \int_Q (\psi_1,\psi_2)\left[\begin{pmatrix} -p_t-b_1\Delta p \\ -q_t-b_2\Delta q \end{pmatrix}+M^T\begin{pmatrix} p \\ q \end{pmatrix}\right]dx\,dt \\
&= \int_Q \left[-\psi_1 p_t-\psi_1 b_1\Delta p+(\psi_1,\psi_2)\,M^T\begin{pmatrix} p \\ q \end{pmatrix}-\psi_2 q_t-\psi_2 b_2\Delta q\right]dx\,dt \\
&= \int_0^T -\left[(p_t,\psi_1)+(q_t,\psi_2)\right]\,dt \\
&\quad+\int_Q\left[b_1\nabla p\nabla\psi_1+b_2\nabla q\nabla\psi_2+(\psi_1,\psi_2)\,M^T\begin{pmatrix} p \\ q \end{pmatrix}\,dx\,dt\right] \\
&= \int_Q (p,q)\,\mathcal{L}\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}\,dx\,dt-\int_\Omega\left[(\psi_1 p)\,(x,T)+(\psi_2 q)\,(x,T)\right]\,dx
\end{aligned}\tag{3.5}$$

The initial and boundary conditions from (3.1) have been used in the last part of (3.5). For the adjoint system, we have the appropriate boundary conditions, zero Neumann conditions, and zero final-time conditions.

THEOREM 3.2. **Existence of Weak Solution**

Given an optimal control h_β and corresponding solution $u_\beta, v_\beta = u(h_\beta), v(h_\beta)$ there exists a weak solution $(p,q)\in V\times V$ satisfying the adjoint system:

$$\begin{aligned}
\mathcal{L}_1^* p &= (u_\beta-z_1)\chi_W+a_1 p-2du_\beta p-h(v_\beta p-v_\beta q) && \textit{in } Q \\
\tfrac{\partial p}{\partial\nu} &= 0 && \textit{on } \partial\Omega\times(0,T)
\end{aligned}\tag{3.6}$$

$$\begin{aligned}
\mathcal{L}_1^* q &= (v_\beta-z_2)\chi_W-a_2 q+h(u_\beta q-u_\beta p) && \textit{in } Q \\
\tfrac{\partial q}{\partial\nu} &= 0 && \textit{on } \partial\Omega\times(0,T)
\end{aligned}\tag{3.7}$$

and transversality conditions

$p(x,T)=0$ *and* $q(x,T)=0$ *when* $t=T, \quad x\in\Omega$.

And furthermore:

$$h_\beta = \min\left(M, \max\left(\frac{1}{\beta}\int_0^T (p-q)u_\beta v_\beta\, dt\,, 0\right)\right). \tag{3.8}$$

Proof :

Let $h_\beta(x)$ be an optimal control (which exists by Theorem 2.2) and (u_β, v_β) be its corresponding state solution. Let $(h_\beta(x)+\varepsilon k)\in U$ for $\varepsilon>0$, and $u^\varepsilon, v^\varepsilon$ be the corresponding weak solution of state system (1.1). Since the adjoint equations are linear, there exists a weak solution p, q satisfying (3.6) and (3.7). We compute the directional derivative of the objective functional $J_\beta(h)$ with respect to h in the direction k at h_β. Since $J_\beta(h_\beta)$ is the minimum value, we have

$$\begin{aligned}
0 &\le \lim_{\varepsilon\to 0^+}\frac{J_\beta(h_\beta+\varepsilon k)-J_\beta(h_\beta)}{\varepsilon}\\
&= \frac{1}{2}\lim_{\varepsilon\to 0^+}\int_W\left[\frac{(u^\varepsilon-z_1)^2-(u_\beta-z_1)^2}{\varepsilon}+\frac{(v^\varepsilon-z_2)^2-(v_\beta-z_2)^2}{\varepsilon}\right]dx\,dt\\
&\quad +\frac{\beta}{2\,\varepsilon}\int_\Omega\left[(h_\beta+\varepsilon k)^2-(h_\beta)^2\right]\,dx\\
&= \int_Q(\psi_1,\psi_2)\begin{pmatrix}(u_\beta-z_1)\chi_W\\ (v_\beta-z_2)\chi_W\end{pmatrix}dx\,dt+\beta\int_\Omega h_\beta k\,dx\\
&= \int_0^T -\left[(p_t,\psi_1)+(q_t,\psi_2)\right]\,dt\\
&\quad +\int_Q\left[b_1\nabla p\nabla\psi_1+b_2\nabla q\nabla\psi_2+(\psi_1,\psi_2)\,M^T\begin{pmatrix}p\\ q\end{pmatrix}\right]dx\,dt+\beta\int_\Omega h_\beta k\,dx\\
&= \int_Q(p,q)\begin{pmatrix}-ku_\beta v_\beta\\ ku_\beta v_\beta\end{pmatrix}dx\,dt+\beta\int_\Omega h_\beta k\,dx
\end{aligned}$$

$$= \int_Q (-pku_\beta v_\beta + qku_\beta v_\beta)\ dx\, dt + \beta \int_\Omega h_\beta k\, dx,$$

where we have used the non-homogeneous terms of the ψ_1, ψ_2 system.

Then we have

$$0 \leq \int_\Omega \left[k \left(\beta h_\beta - \int_0^T (p-q) u_\beta v_\beta \, dt \right) \right] \, dx \tag{3.9}$$

We know that our control $h_\beta(x)$ satisfies $0 \leq h_\beta \leq M$.

Case 1: On the set where $0 < h_\beta < M$, the variation function k can be of any sign.

Suppose $\beta h_\beta - \displaystyle\int_0^T (p-q) u_\beta v_\beta \, dt > 0$ on some set with positive measure, and take $k < 0$ with support on this set, then

$0 > \displaystyle\int_\Omega \left[k \left(\beta h_\beta - \int_0^T (p-q) u_\beta v_\beta \, dt \right) \right] \, dx,$ which is clearly a contradiction to (3.9).

Suppose $\beta h_\beta - \displaystyle\int_0^T (p-q) u_\beta v_\beta \, dt < 0$ on some set with positive measure and $k > 0$ with support on this set, then

$0 > \displaystyle\int_\Omega \left[k \left(\beta h_\beta - \int_0^T (p-q) u_\beta v_\beta \, dt \right) \right] \, dx,$ which is clearly a contradiction to (3.9).

Thus, we have

$$\beta h_\beta - \textstyle\int_0^T (p-q) u_\beta v_\beta \, dt = 0$$

and

$$0 < h_\beta = \frac{1}{\beta} \int_0^T (p-q) u_\beta v_\beta \, dt < M$$

Case 2 : On the set where $h_\beta = 0$, the variation function must satisfy $k \geq 0$, which implies

$$0 = h_\beta \geq \frac{1}{\beta} \int_0^T (p-q) u_\beta v_\beta \, dt$$

Case 3 : On the set where $h_\beta = M$, the variation function must satisfy $k \le 0$, which implies

$$\beta h_\beta - \int_0^T (p-q)u_\beta v_\beta\, dt \le 0$$

and

$$M = h_\beta \le \frac{1}{\beta}\int_0^T (p-q)u_\beta v_\beta\, dt.$$

Hence :

$$h_\beta = \begin{cases} \frac{1}{\beta}\int_0^T (p-q)u_\beta v_\beta\, dt & \text{if} \quad 0 < \frac{1}{\beta}\int_0^T (p-q)u_\beta v_\beta\, dt < M \\ 0 & \text{if} \quad \frac{1}{\beta}\int_0^T (p-q)u_\beta v_\beta\, dt \le 0 \\ M & \text{if} \quad \frac{1}{\beta}\int_0^T (p-q)u_\beta v_\beta\, dt \ge M \end{cases}$$

Using compact notation, the above result can be rewritten as:

$$h_\beta = \min\left(M, \max\left(\frac{1}{\beta}\int_0^T (p-q)u_\beta v_\beta\, dt\,, 0\right)\right). \quad \Box$$

Substituting h_β in system (1.1) and in adjoint equations (3.6) and (3.7), we obtain the following optimality system (OS):

$$\begin{aligned} \mathcal{L}_1 u &= (a_1 - du)u - \min\left(M, \max\left(\tfrac{1}{\beta}\textstyle\int_0^T (p-q)uv\, dt\,, 0\right)\right) uv && \text{in } Q \\ \mathcal{L}_2 v &= -a_2 v + \min\left(M, \max\left(\tfrac{1}{\beta}\textstyle\int_0^T (p-q)uv\, dt\,, 0\right)\right) uv && \text{in } Q \\ \mathcal{L}_1^* p &= (u - z_1)\chi_W + a_1 p - 2dup \\ &\quad - \min\left(M, \max\left(\tfrac{1}{\beta}\textstyle\int_0^T (p-q)uv\, dt\,, 0\right)\right)(vp - vq) && \text{in } Q \\ \mathcal{L}_2^* q &= (v - z_2)\chi_W - a_2 q \\ &\quad + \min\left(M, \max\left(\tfrac{1}{\beta}\textstyle\int_0^T (p-q)uv\, dt\,, 0\right)\right)(uq - up) && \text{in } Q \\ & u(x,0) = u_0(x), \quad v(x,0) = v_0(x) \qquad x \in \Omega \\ & p(x,T) = 0, \quad q(x,T) = 0 \qquad x \in \Omega \\ & \tfrac{\partial u}{\partial \nu} = 0 = \tfrac{\partial v}{\partial \nu} \qquad \text{on } \partial\Omega \times (0,T) \\ & \tfrac{\partial p}{\partial \nu} = 0 = \tfrac{\partial q}{\partial \nu} \qquad \text{on } \partial\Omega \times (0,T). \end{aligned} \tag{3.10}$$

The weak solution of the optimality system exists by Theorems 2.2 and 3.2. For small T, we now prove uniqueness of weak solutions of the optimality system, which gives the characterization (3.8) of the unique optimal control in terms of the solutions of the OS.

THEOREM 3.3. *For T sufficiently small, weak solutions of the optimality system are unique.*

Proof: Suppose (u, v, p, q) and $(\bar{u}, \bar{v}, \bar{p}, \bar{q})$ are two weak solutions. We do a change of variables:

$$
\begin{aligned}
u &= e^{\lambda t} w, \quad \bar{u} = e^{\lambda t}\bar{w}, \quad v = e^{\lambda t} z, \quad \bar{v} = e^{\lambda t}\bar{z} \\
p &= e^{-\lambda t} y, \quad \bar{p} = e^{-\lambda t}\bar{y}, \quad q = e^{-\lambda t}\xi, \quad \bar{q} = e^{-\lambda t}\bar{\xi},
\end{aligned}
$$

where $\lambda > 0$ is to be chosen. Denote

$$
\begin{aligned}
h &= \min\left(M, \max\left(\frac{1}{\beta}\int_0^T (p-q)uv\,dt\,, 0\right)\right) \\
\bar{h} &= \min\left(M, \max\left(\frac{1}{\beta}\int_0^T (\bar{p}-\bar{q})\bar{u}\bar{v}\,dt\,, 0\right)\right),
\end{aligned}
$$

We subtract the PDEs corresponding to w and $\bar{w}$, z and $\bar{z}$, y and $\bar{y}$, ξ and $\bar{\xi}$. The weak formulation of $w - \bar{w}$ is illustrated:

$$
\begin{aligned}
&\int_Q (w-\bar{w})_t \phi\,dx\,dt + \int_Q b_1 \nabla(w-\bar{w})\phi\,dx\,dt + \lambda\int_Q (w-\bar{w})\phi\,dx\,dt \\
&= \int_Q a_1(w-\bar{w})\phi\,dx\,dt - \int_Q de^{\lambda t}(w^2-\bar{w}^2)\phi\,dx\,dt - \int_Q e^{\lambda t}(hwz - \bar{h}\bar{w}\bar{z})\phi\,dx\,dt,
\end{aligned}
$$

where $\phi \in V$ (a test function).

By standard estimation techniques and using the test functions $(w-\bar{w})$, $(z-\bar{z})$, $(y-\bar{y})$, $(\xi-\bar{\xi})$ in the PDEs, we obtain:

$$
\begin{aligned}
&\frac{1}{2}\int_\Omega \{[w(x,T)-\bar{w}(x,T)]^2 + [z(x,T)-\bar{z}(x,T)]^2 + [y(x,0)-\bar{y}(x,0)]^2 \\
&+[\xi(x,0)-\bar{\xi}(x,0)]^2\}\,dx \\
&+\int_Q [b_1(|\nabla(w-\bar{w})|)^2 + b_2(|\nabla(z-\bar{z})|)^2 + b_1(|\nabla(y-\bar{y})|)^2 + b_2(|\nabla(\xi-\bar{\xi})|)^2]\,dx\,dt
\end{aligned}
$$

$$
\begin{aligned}
&+\lambda \int_Q [(w-\bar{w})^2 + (z-\bar{z})^2 + (y-\bar{y})^2 + (\xi-\bar{\xi})^2]\,dx\,dt \\
= \;& \int_Q [a_1(w-\bar{w})^2 - a_2(z-\bar{z})^2 + a_1(y-\bar{y})^2 - a_2(\xi-\bar{\xi})^2]\,dx\,dt \\
&- \int_Q de^{\lambda t}(w^2-\bar{w}^2)(w-\bar{w})\,dx\,dt \\
&+ \int_Q e^{2\lambda t}\left[(w-\bar{w})(y-\bar{y}) + (z-\bar{z})(\xi-\bar{\xi})\right]\chi_W\,dx\,dt \\
&- \int_Q e^{\lambda t}\left[(hwz-\bar{h}\bar{w}\bar{z})(w-\bar{w}) - (hwz-\bar{h}\bar{w}\bar{z})(z-\bar{z})\right]\,dx\,dt \\
&- \int_Q \left\{2de^{\lambda t}(wy-\bar{w}\bar{y}) - e^{\lambda t}\left[hz(y+\xi) - \bar{h}\bar{z}(\bar{y}+\bar{\xi})\right]\right\}(y-\bar{y})\,dx\,dt \\
&+ \int_Q e^{\lambda t}\left[hw(\xi-y) - \bar{h}\bar{w}(\bar{\xi}-\bar{y})\right](\xi-\bar{\xi})\,dx\,dt \\
= \;& \int_Q [a_1(w-\bar{w})^2 - a_2(z-\bar{z})^2 + a_1(y-\bar{y})^2 - a_2(\xi-\bar{\xi})^2]\,dx\,dt \\
&- \int_Q de^{\lambda t}(w-\bar{w})^2(w+\bar{w})\,dx\,dt \\
&+ \int_Q e^{2\lambda t}\left[(w-\bar{w})(y-\bar{y}) + (z-\bar{z})(\xi-\bar{\xi})\right]\chi_W\,dx\,dt \\
&- \int_Q e^{\lambda t}\left[(hw(z-\bar{z})(w-\bar{w}) + h\bar{z}(w-\bar{w})^2 + \bar{z}\bar{w}(h-\bar{h})(w-\bar{w})\right] \\
&+ \int_Q e^{\lambda t}\left[hw(z-\bar{z})^2 + h\bar{z}(w-\bar{w})(z-\bar{z}) + \bar{z}\bar{w}(h-\bar{h})(z-\bar{z})\right]\,dx\,dt \\
&- \int_Q 2de^{\lambda t}\left[y(w-\bar{w})(y-\bar{y}) + \bar{w}(y-\bar{y})^2\right]\,dx\,dt \\
&+ \int_Q e^{\lambda t}\left[hz(y-\bar{y})^2 + h\bar{y}(z-\bar{z})(y-\bar{y}) + \bar{z}\bar{y}(h-\bar{h})(y-\bar{y})\right]\,dx\,dt \\
&+ \int_Q e^{\lambda t}\left[hz(\xi-\bar{\xi})(y-\bar{y}) + h\bar{\xi}(z-\bar{z})(y-\bar{y}) + \bar{z}\bar{\xi}(h-\bar{h})(y-\bar{y})\right]\,dx\,dt \\
&+ \int_Q e^{\lambda t}\left[hw(\xi-\bar{\xi})^2 + h\bar{\xi}(w-\bar{w})(\xi-\bar{\xi}) + \bar{w}\bar{\xi}(h-\bar{h})(\xi-\bar{\xi})\right]\,dx\,dt \\
&- \int_Q e^{\lambda t}\left[hw(y-\bar{y})(\xi-\bar{\xi}) + h\bar{y}(w-\bar{w})(\xi-\bar{\xi}) + \bar{w}\bar{y}(h-\bar{h})(\xi-\bar{\xi})\right]\,dx\,dt
\end{aligned}
$$

Now we use the the fact that w, z, y and ξ are L^∞ bounded. We also use Cauchy inequality in the RHS terms. The specific terms are combined using:

$$
\begin{aligned}
\int_Q de^{\lambda t}(w-\bar{w})^2(w+\bar{w})\,dx\,dt &\leq c_1 e^{\lambda T}\int_Q (w-\bar{w})^2\,dx\,dt \\
\int_Q \left|e^{2\lambda t}(w-\bar{w})(y-\bar{y})\right|\,dx\,dt &\leq c_2 e^{\lambda 2T}\int_Q [(w-\bar{w})^2+(y-\bar{y})^2]\,dx\,dt \\
\int_Q \left|e^{\lambda t}\bar{z}\bar{w}(h-\bar{h})(w-\bar{w})\right|\,dx\,dt &\leq c_3 e^{\lambda T}\int_Q [(w-\bar{w})^2+(h-\bar{h})^2]\,dx\,dt.
\end{aligned}
$$

Also, we estimate

$$
\begin{aligned}
|h-\bar{h}| &\leq \left|\frac{1}{\beta}\int_0^T e^{\lambda t}[(y-\xi)wz-(\bar{y}-\bar{\xi})\bar{w}\bar{z}]\,dt\right| \\
&\leq \left|\frac{1}{\beta}\int_0^T e^{\lambda t}[wz(\xi-\bar{\xi})+w\bar{\xi}(z-\bar{z})+\bar{z}\bar{\xi}(w-\bar{w})\right. \\
&\qquad \left. -wz(y-\bar{y})-w\bar{y}(z-\bar{z})-\bar{z}\bar{y}(w-\bar{w})]\,dt\,\right|
\end{aligned}
$$

Thus, we obtain

$$
|h-\bar{h}|^2 \leq c_3 e^{\lambda 2T}\int_0^T \left[(w-\bar{w})^2+(z-\bar{z})^2+(y-\bar{y})^2+(\xi-\bar{\xi})^2\right]\,dt.
$$

Hence

$$
\begin{aligned}
&\frac{1}{2}\int_\Omega \{[w(x,T)-\bar{w}(x,T)]^2+[z(x,T)-\bar{z}(x,T)]^2+[y(x,0)-\bar{y}(x,0)]^2 \\
&\qquad +[\xi(x,0)-\bar{\xi}(x,0)]^2\}\,dx \\
&+\int_Q [b_1|\nabla(w-\bar{w})|^2+b_2|\nabla(z-\bar{z})|^2+b_1|\nabla(y-\bar{y})|^2+b_2|\nabla(\xi-\bar{\xi})|^2]\,dx\,dt \\
&+(\lambda-C_1-C_2e^{\lambda 2T}-C_3Te^{\lambda 2T})\int_Q [(w-\bar{w})^2+(z-\bar{z})^2+(y-\bar{y})^2+b_2(\xi-\bar{\xi})^2]\,dx\,dt \\
&\leq 0
\end{aligned}
$$

where C_1, C_2, C_3 depend on the coefficients, the L^∞ bounds of w, z, y and ξ, and C_3T term comes from the $h-\bar{h}$ term.

If we choose λ such that $\lambda > C_1 + C_2 + C_3$ and T small such that $\lambda - C_1 - C_2 e^{\lambda 2T} - C_3 T e^{\lambda 2T} > 0$ then $w = \bar{w}, z = \bar{z}, y = \bar{y}$ and $\xi = \bar{\xi}$. Therefore we conclude $u = \bar{u}, v = \bar{v}, p = \bar{p}$ and $q = \bar{q}$. $\square$

Remark: The T sufficiently small condition in the uniqueness result is due to opposite time orientations of the state and adjoint system. The approximate coefficient,
$h_\beta = \min\left(M, \max\left(\frac{1}{\beta}\int_0^T (p-q)u_\beta v_\beta \, dt, 0\right)\right)$ is characterized in terms of (u_β, v_β, p, q), the unique solution of the optimality system for sufficiently small T. This smallness condition is usual in these parabolic optimality systems [**10, 16**].

4. Identification Problem

The objective functional (1.3) is an approximation in the sense that the second term is artificial as always in the Tikhonov regularization. We now use a sequence of optimal controls as $\beta \to 0$ to identify h from observations z_1 and z_2. We do not assume that (z_1, z_2) is in the range of the map

$$h \in U \to (u(h), v(h))|_W \tag{4.1}$$

This situation could occur from inaccurate observations.

THEOREM 4.1. *There exist*

(i) *a sequence* $\beta_n \to 0$

(ii) *a corresponding sequence of optimal controls,* h_{β_n}*, for the functionals* $J_{\beta_n}(h)$*,*

(iii) $h^* \in U$ *and*

(iv) $u^* = u(h^*), v^* = v(h^*)$ *(solution of (1.1)), such that*

$$\begin{aligned} h_{\beta_n} &\rightharpoonup h^* \quad \textit{weakly in } L^2(Q) \\ u_{\beta_n} = u(h_\beta) &\rightharpoonup u^* \quad \textit{weakly in } L^2(0,T;H^1(\Omega)) \\ v_{\beta_n} = v(h_\beta) &\rightharpoonup v^* \quad \textit{weakly in } L^2(0,T;H^1(\Omega)) \end{aligned}$$

and

$$\int_W \left[(u^* - z_1)^2 + (v^* - z_2)^2\right] \, dx\, dt = \inf_{h \in U} \int_W \left[(u(h) - z_1)^2 + (v(h) - z_2)^2\right] \, dx\, dt. \tag{4.2}$$

Note: The limits u^*, v^* can be interpreted as a (not necessarily unique) projection of (z_1, z_2) onto the range of the map (4.1).

Proof: We note that the a *priori* estimates (2.20) are independent of β. Thus we can find a sequence $\beta_n \to 0$ and $h^* \in U, u^*, v^* \in L^2(0, T; H^1(\Omega))$ such that

$$J_{\beta_n}(h_{\beta_n}) = \inf_{h \in U} J_{\beta_n}(h)$$

and

$$\begin{array}{rll} h_{\beta_n} & \rightharpoonup h^* & \text{weakly in } L^2(Q) \\ u_{\beta_n} & \rightharpoonup u^* & \text{weakly in } L^2(0, T; H^1(\Omega)) \\ v_{\beta_n} & \rightharpoonup v^* & \text{weakly in } L^2(0, T; H^1(\Omega)) \\ u_{\beta_n} & \to u^* & \text{strongly in } L^2(Q) \\ v_{\beta_n} & \to v^* & \text{strongly in } L^2(Q) \\ (u_{\beta_n})_t & \rightharpoonup (u^*)_t & \text{weakly in } L^2(0, T; (H^1(\Omega))^*) \\ (v_{\beta_n})_t & \rightharpoonup (v^*)_t & \text{weakly in } L^2(0, T; (H^1(\Omega))^*) \end{array}$$

where $(u^*, v^*) = (u(h^*), v(h^*))$ is the weak solution of (1.1).

Now for any $\bar{h} \in U, \bar{u} = u(\bar{h}), \bar{v} = v(\bar{h})$, we have

$$J_{\beta_n}(h_{\beta_n}) \leq J_{\beta_n}(\bar{h}), \tag{4.3}$$

i.e.,

$$\begin{aligned} &\frac{1}{2} \int_W [(u_{\beta_n} - z_1)^2 + (v_{\beta_n} - z_2)^2] \, dx\, dt \; + \frac{\beta_n}{2} \int_\Omega h_{\beta_n}(x)^2 \, dx \\ &\leq \frac{1}{2} \int_W [(\bar{u} - z_1)^2 + (\bar{v} - z_2)^2] \, dx\, dt \; + \frac{\beta_n}{2} \int_\Omega \bar{h}(x)^2 \, dx. \end{aligned}$$

Letting $\beta_n \to 0$ in (4.3) gives

$$\int_W \left[(u^* - z_1)^2 + (v^* - z_2)^2\right] \, dx\, dt \leq \int_W \left[(\bar{u} - z_1)^2 + (\bar{v} - z_2)^2\right] \, dx\, dt,$$

which yields the desired result (4.2). $\square$

Remark: If (z_1, z_2) were in the range of map (4.1) then there exists $h_0 \in U$ such that $u(h_0) = z_1, v(h_0) = z_2$ and the RHS of (4.2) would vanish. If the solution to the inverse problem were unique, the identified coefficient h^* would equal h_0, the solution of the inverse problem.

5. Numerical Example

In this section we present a numerical example for a test problem. First we illustrate an iterative algorithm and its explicit finite difference discretization for solving the optimality system. Our optimality system (3.10) consists of a pair of state equations coupled with a pair of adjoint equations and the characterization (3.8) for h_β. In each iteration, the state system is solved forward in time while the adjoint system is solved backward in time. The convergence of such forward-backward iteration schemes was proven by Hackbush [**11**].

The following steps are followed:

Step 1: Choose (u_0, v_0), (p_T, q_T) and h.

Step 2: Compute (u, v) by solving the following system forward in time

$$\begin{aligned} u_t - b_1 \Delta u &= (a_1 - du)u - h(x)uv && \text{in } Q = \Omega \times (0,T) \\ v_t - b_2 \Delta v &= -a_2 v + h(x)uv && \text{in } Q = \Omega \times (0,T) \end{aligned} \tag{5.1}$$

$$\begin{aligned} u(x,0) &= u_0(x), \quad v(x,0) = v_0(x) && \text{for } x \in \Omega \\ \frac{\partial u}{\partial \nu} &= 0, \quad \frac{\partial v}{\partial \nu} = 0 && \text{on } \partial\Omega \times (0,T). \end{aligned}$$

Step 3: Compute (p, q) by solving the following system backward in time

$$\begin{aligned} -p_t - b_1 \Delta p &= a_1 p - 2dup - h(vp - vq) + (u - z_1)\chi_W && \text{in } Q \\ -q_t - b_2 \Delta q &= -a_2 q + h(uq - up) + (v - z_2)\chi_W \end{aligned} \tag{5.2}$$

$$\frac{\partial p}{\partial \nu} = 0, \quad \frac{\partial q}{\partial \nu} = 0 \qquad \text{on } \partial\Omega \times (0,T)$$
$$p(x,T) = 0, \quad q(x,T) = 0 \quad \text{when} \quad t = T, \quad x \in \Omega,$$

where χ_W is the characteristic function of the set W.

Step 4: Update the control using u, v, p, q from steps 2 and 3:

$$h_{new} = \min\left(M, \max\left(\frac{1}{\beta}\int_0^T (p-q)uv\,dt\,, 0\right)\right). \tag{5.3}$$

Step 5: Continue the loop until we get the convergence results with desired accuracy, i.e., the state and adjoint values at the current iteration are very close to the corresponding values at the previous iteration.

Clearly, at each iteration the system of equations (5.1) and (5.2) needs to solved, which are coupled by (5.3). The equations are discretized using the following [**1**]:

$$u(x_i,t_j)_t \approx \frac{1}{\Delta t}\left[u(x_i,t_{j+1}) - u(x_i,t_j)\right]$$
$$u(x_i,t_j)_{xx} \approx \frac{1}{\Delta x^2}\left[u(x_{i+1},t_j) - 2u(x_i,t_j) + u(x_{i-1},t_j)\right]$$

The time integration in h_{new} given by (5.3) is approximated by the composite Simpson's rule [**27**].

We illustrate our procedure on the following one spatial dimension problem:

$$u_t - b_1 u_{xx} = (a_1 - du)u - h(x)uv \qquad \text{in } (0,1)\times(0,1)$$
$$v_t - b_2 v_{xx} = -a_2 v + h(x)uv \qquad \text{in } (0,1)\times(0,1)$$
$$u(x,0) = u_0(x), \quad v(x,0) = v_0(x) \qquad \text{for } x \in (0,1)$$
$$\frac{\partial u}{\partial \nu} = 0, \qquad \frac{\partial v}{\partial \nu} = 0 \qquad \text{for } t \in (0,1),\ x = 0 \text{ and } x = 1.$$

where

$$h(x) = 0.25 + x$$
$$u(x,t) = e^{-t}(2 + \cos \pi x)$$
$$v(x,t) = e^{-2t}(2 - \cos \pi x)$$

$$
\begin{aligned}
a_1 &= \frac{[-2-\cos\pi x(1-b_1\pi^2)+h(x)e^{-2t}(4-\cos^2\pi x)+de^{-t}(2+\cos\pi x)^2]}{2+\cos\pi x} \\
a_2 &= \frac{[-4+\cos\pi x(2-b_2\pi^2)-h(x)e^{-t}(4-\cos^2\pi x)]}{-2+\cos\pi x}
\end{aligned}
$$

and we choose $b_1 = 1/\pi^2, b_2 = 2/\pi^2, d = 2e^4$ so that $a_1, a_2 \geq 0$. The coefficient β and M in (5.3) are chosen 0.14 and 500.

The exact observations (z_1, z_2) in the objective functional ($J_\beta(h)$) are chosen to be $u(x,t), v(x,t)$ in the first iteration.

Our aim is to recover the coefficient $h(x)$. The exact and approximated solution are plotted in the Figure 9.

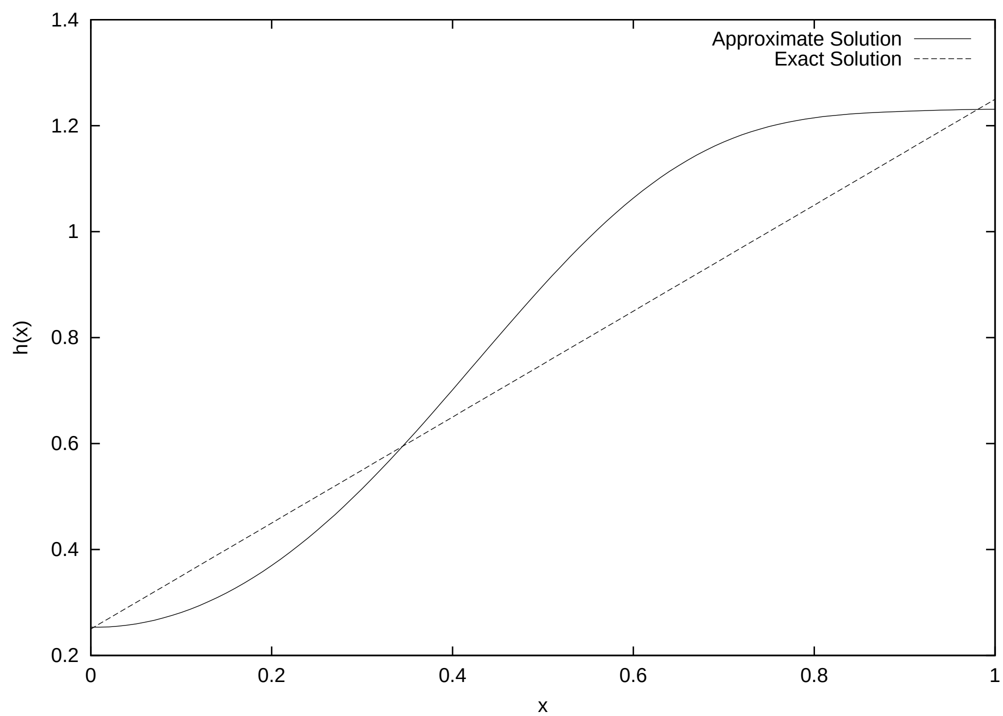

FIGURE 9. *Exact* $h(x)$ *and Computed* $h_\beta(x)$

Bibliography

[1] Ames, W. F., *Numerical Methods for Partial Differential Equations*, Barnes and Noble, 1969.

[2] Anikonov, Y. E. and Lorenzi, A., *Existence Results for Parabolic Identification Problems*, J. Inv. Ill-Posed Problems, Vol. 2, 3(1994), 181-210.

[3] Banks, H. T. and Kunish K., *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston 1989.

[4] Chavent, G., *Identification and System Parameter Estimation*. IFAC, York, UK, 1985.

[5] Chavent, G. and Kunish, K., *Convergence of Tikhonov Regularization for Constrained Ill-posed Inverse Problem*, Inverse Problems, 10(1994), 63-78.

[6] Choulli, M. and Yamamoto, M., *Generic well-posedness of an inverse parabolic problem–the Hölder-space approach*, Inverse Problems, 13(1996), 195-205.

[7] Elayyan, A., *Uniqueness for a One-Dimensional Inverse Parabolic Problem*, Inverse Problems, 15(1999), 1283-1294.

[8] Engl, H. W. and Zou, J., *A New Approach to Convergence Rate Analysis of Tikhonov Regularization for Parameter Identification in Heat Conduction*, Inverse Problems, 16(2000), 1907-1923.

[9] Evans, L.C., *Partial Differential Equations*, American Mathematical Society, Providence, 1998.

[10] Fister, R., *Optimal Control of Harvesting in a Predator-Prey Parabolic System*, Houston J. Math. 23(1997), No. 2, 341-355.

[11] Hackbush, W., K., *A Numerical Method for Solving Parabolic Equations with Opposite Orientation*, Computing, 20(1978), 229-240.

[12] Imanuvilov, O. Y. and Yamamoto, M., *Lipschitz Stability in Inverse Parabolic Problem by the Carleman Estimate*, Inverse Problems, 14(1998), 1229-1245.

[13] Isakov, V., *Uniqueness and Stability in Inverse parabolic Problems*, Inverse Problems in Diffusion Process, Proceedings of the GAMM-SIAM Symposium, edited by H. W. Engl and W. Rundell, 21-41.

[14] Krylov, N.V., *Nonlinear Elliptic and Parabolic Equations of the Second Order*, Mathematics and its Applications, Reidel, 1987.

[15] Lenhart, S., Feng, X., Protopopescu, V., Rachele, L. and Sutton, B., *Identification Problem for the Wave Equation with Neumann Data Input and Dirichlet Data Observations*, Preprint 2001.

[16] Lenhart, S., Liang, M. and Protopopescu, V., *Optimal control of Boundary Habitat Hostility for Interacting Species*, Math. Meth. Appl. Sci. **22**, 1061-1077 (1999).

[17] Lenhart, S., Liang, M., and Protopopescu, V., *Identification Problem for a Wave Equation via Optimal Control*, Control of Distributed Parameter and Stochastic Systems, Kluner Academic Publisher, 1998, 79-84.

[18] Lenhart, S., Protopopescu, V. and Yong, J., *Solving Inverse Problems of Identification Type by Optimal Control Methods*, Proceedings of the International Conference on Advances on Nonlinear Dynamics near the Millennium, ANDM'97, 1997, 87-94.

[19] Lenhart, S., Protopopescu, V. and Yong, J., *Identification of Boundary Shape and Reflectivity in a Wave Equation by Optimal Control Techniques*, Direct and Inverse Problems of Mathematical Physics, Kulwer Academic Publisher, 2000, 251-266.

[20] Lenhart, S., Protopopescu, V. and Yong, J., *Optimal Control of a Reflecting Boundary Coefficient in an Acoustic Wave Equation*, Applicable Analysis, 68(1998), 179-194.

[21] Leung, A. W., *Optimal Harvesting-Coefficient Control of Steady-Sate Prey-Predator Diffusive Volterra-Lotka Systems*, Appl. Math. Optim., 31(1995), no 2, 219-241.

[22] Leung, A. W., and Stojanovic S., *Optimal Control for Elliptic Volterra-Lotka Type Equations*, J. Math. Anal. Appl., 173 (1993), no. 2, 603–619.

[23] Lions, J. L., *Optimal Control of Systems Governed by Partial Differential Equations*, Springer-Verlag, New York, 1971.

[24] Murray, J. D. , *Mathematical Biology*, Springer-Verlag, Berlin, 1989.

[25] Protter, M.H. and Weinberger, H.F. , *Maximum Principle in Differential Equations*, Prentice Hall Inc, Englewood Cliffs, 1967.

[26] Simon, J., *Compact sets in the space* $L^2(0,T;B)$, Annali di Matematica Pura et Applicata CXLVI (1987), 65-96.

[27] Stoer, J. and Bulrish, R., *Introduction to Numerical Analysis*, Springer Verlag, New York, 1993.

[28] Tikhonov, A. N. and Arsenin, V. Y., *Solutions of Ill-posed Problems*, John Wiley, New York 1997.

[29] Yamamoto, M., *Conditional Stability in Determination of Force Terms of Heat Equations in a Rectangle*, Math. Comput. Modeling, Vol. 18, 1(1993), 79-83.

Part IV

Control of the Convective Velocity Coefficient in a Parabolic Problem

1. Introduction

We consider optimal control of the coefficient of the first spatial derivative in a parabolic equation. We treat the case of one spatial dimension and the control depends on space and time. Control of such a coefficient requires more regularity on the control set than controlling the coefficient of a zero$^{\text{th}}$ order term, and it leads to a partial differential equation (PDE) characterization of an optimal control instead of the usual explicit characterization.

We consider the parabolic equation (in non divergent form)

$$\begin{aligned} u_t(x,t) &= au_{xx} - h(x,t)u_x + f \quad \text{in } Q = (0,L) \times (0,T) \\ u(x,0) &= u_0(x) \qquad\qquad\qquad 0 < x < L,\ t = 0, \end{aligned} \tag{1.1}$$

with the boundary conditions

$$\begin{aligned} u(0,t) = 0 &\qquad \text{for all} \quad t \in (0,T) \\ u_x(L,t) = 0 &\qquad \text{for all} \quad t \in (0,T). \end{aligned}$$

The convective velocity coefficient $h(x,t)$ is taken to be the control.

The objective functional is given by

$$J(h) = \alpha \int_0^T \int_0^L u(x,t)\,dx\,dt + \frac{\beta}{2} \int_0^T \int_0^L \left[h^2(x,t) + h_t^2(x,t) + h_x^2(x,t)\right]\,dx\,dt,$$

where α and β are positive constants known as weighting functions.

Our goal is to minimize the objective functional:

$$J(h^*) = \min_{h \in U} J(h)$$

with control set $U = H^1(Q)$.

The example is designed to illustrate this technique with more regular controls and to generalize two recent flow applications. Objective functionals with more general differentiable dependence on u can also be treated. Handagama and Lenhart [**3**] considered an optimal control of the time dependent flow rate into a gas phase bioreactor, where the degradation of a contaminant is through the metabolism of a bacteria in the bioreactor. Lenhart [**4**] studied an optimal control of a parabolic differential equation, modeling one dimensional fluid through a soil-packed tube in which a contaminant is initially distributed. She considered the convective velocity coefficient as a function of time only. See Chawla and Lenhart [**1**] for a similar application involving bioreactor flow.

In section 2, the problem is formulated and the existence of an optimal control is proven. In section 3, the optimality system is derived by differentiating the objective functional with respect to the control (in a directional derivative) and differentiating the function mapping the control to the state. We obtain the characterization of an optimal control through an elliptic partial differential equation. Thus the optimality system will consist of two parabolic equations coupled with an elliptic equation.

2. Existence of Optimal Control

The solution of the state equation is a weak solution in

$$W = L^2(0, T; V)$$

where $V = H^1_{\{0\}}(0, L)$. (Subscript $\{0\}$ means the functions vanish at $x = 0$.)

DEFINITION 2.1. *A function $u \in W$ with $u_t \in L^2(0, T; V^*)$ is a solution of*

$$\begin{aligned} u_t(x,t) &= au_{xx} - h(x,t)u_x + f && \text{in } Q = (0,L) \times (0,T) \\ u(0,t) &= 0 && \text{on } (0,T), x = 0 \\ u_x(L,t) &= 0 && \text{on } (0,T), x = L \\ u(x,0) &= u_0(x) && \text{on } (0,L), t = 0 \end{aligned} \tag{2.1}$$

if

$$\langle u'(t), v \rangle + \int_0^L (au_x v_x + hu_x v - fv)\, dx = 0 \tag{2.2}$$

for all $v \in V$ and a. e. time $0 \leq t \leq T$, and $u(x,0) = u_0(x)$, where the brackets $<,>$ in the integrand denote the duality between V and V^*.

Note: Since $u \in C(0,T;L^2(0,L))$ from Evans [**2**], the initial condition of (2.1) makes sense.

The following assumptions are made:

(a) $a,\ \alpha,\ \beta$ are positive constants,

(b) $u_0 \in H^1(0,L)$, with compact support,

(c) $f \in L^\infty(Q)$,

For $h \in U$, the existence of a unique solution $u = u(h)$ in W satisfying (2.1) in the sense (2.2) follows from a standard result in [**2**].

THEOREM 2.1. *There exists an optimal control $h^* \in U$ minimizing the objective functional $J(h)$.*

PROOF. Choose a minimizing sequence $\{h^n\}, n = 1, 2, \ldots$, such that

$$\lim_{n\to\infty} J(h^n) = \inf_{h\in U} J(h).$$

Let $u^n = u(h^n)$ be the corresponding state solutions to (2.1). To obtain the necessary *a priori* estimates on our minimizing sequence and associated states, we need a change of variable $w = e^{-\lambda t}u$, where λ is a positive constant to be chosen. Thus equation (1.1) becomes

$$w_t + \lambda w = a w_{xx} - h(x,t) w_x + e^{-\lambda t} f.$$

We use the weak definition on the above equation and integrate in time to obtain

$$\int_0^T \langle w_t^n, w^n \rangle \, dt + \int_0^T \int_0^L \left(a w_x^n w_x^n + h^n w_x^n w^n + \lambda (w^n)^2 - e^{-\lambda t} w^n f \right) dx\, dt = 0$$

and then

$$\frac{1}{2}\int_0^L (w^n)^2(x,T)\, dx + a\int_0^T \int_0^L (w_x^n)^2 \, dx\, dt + \lambda \int_0^T \int_0^L (w^n)^2 \, dx\, dt$$

$$= \frac{1}{2}\int_0^L u_0^2\,dx - \int_0^T\int_0^L h^n w_x^n w^n\,dx\,dt + \int_0^T\int_0^L e^{-\lambda t} w^n f\,dx\,dt.$$

Note: $\|h^n\|_{H^1(Q)} \le C$ by the form of the objective functional and $h^n \in H^1(Q) \subset C^{0,1/2}(\bar{Q})$ [**6**]. Thus there exists a constant M such that $\|h^n\|_{L^\infty} \le M$.

Now we apply Cauchy's inequality in the right hand side to obtain

$$\begin{aligned}
&\frac{1}{2}\int_0^L (w^n)^2(x,T)\,dx + a\int_0^T\int_0^L (w_x^n)^2\,dx\,dt + \lambda\int_0^T\int_0^L (w^n)^2\,dx\,dt \\
&\le \frac{1}{2}\int_0^L u_0^2\,dx + \frac{a}{2}\int_0^T\int_0^L (w_x^n)^2\,dx\,dt + \frac{M^2}{2a}\int_0^T\int_0^L (w^n)^2\,dx\,dt \\
&\quad + \frac{1}{4}\int_0^T\int_0^L (w^n)^2\,dx\,dt + \int_0^T\int_0^L f^2\,dx\,dt.
\end{aligned}$$

Choose λ such that $\lambda > 1 + \dfrac{M^2}{a}$. Thus we obtain:

$$\begin{aligned}
&\frac{1}{2}\int_0^L (w^n)^2(x,T)\,dx + a\int_0^T\int_0^L (w_x^n)^2\,dx\,dt + \frac{M^2}{2a}\int_0^T\int_0^L (w^n)^2\,dx\,dt \\
&\le \frac{1}{2}\int_0^L u_0^2\,dx + \int_0^T\int_0^L f^2\,dx\,dt.
\end{aligned} \tag{2.3}$$

By (2.3) we can conclude that $\|w^n\|_W$ are uniformly bounded independent of n. Thus $\|u^n\|_W$ are uniformly bounded independent of n. Also using (1.1) and the above bounds, we have uniform bounds on $\|u_t^n\|_{L^2(0,T;V^*)}$.

Using the above bounds, we can extract a subsequence satisfying the following convergence properties.

$$\begin{aligned}
u^n &\rightharpoonup u^* \quad \text{weakly in } W \\
u_t^n &\rightharpoonup u_t^* \quad \text{weakly in } L^2(0,T;V^*) \\
h^n &\to h^* \quad \text{strongly in } H^1(Q), \\
h_x^n &\rightharpoonup h_x^* \quad \text{weakly in } L^2(Q), \\
h_t^n &\rightharpoonup h_t^* \quad \text{weakly in } L^2(Q),
\end{aligned} \tag{2.4}$$

The strong convergence of h^n in L^2 makes sense as $H^1(Q) \subset\subset L^2(Q)$.

Now, we need to show

$$u^* = u(h^*), \quad \text{i.e. } u^* \text{ is the state solution associated with } h^*.$$

Consider the weak form of the PDE (integrated in time) satisfied by u^n (for $v \in V$):

$$\int_0^T \langle u_t^n, v\rangle \, dt + a\int_0^T \int_0^L u_x^n v_x \, dx\, dt + \int_0^T \int_0^L h^n u_x^n v \, dx\, dt - \int_0^T \int_0^L fv \, dx\, dt = 0.$$

The convergence of the third term uses the fact that $h^n \in H^1(Q) \subset C^{0,1/2}(\bar{Q})$. We have

$$\begin{aligned}
&\left|\int_0^T \int_0^L h^n u_x^n v - h u_x v \, dx\, dt\right| \\
&\le \int_0^T \int_0^L |(h^n - h)v||u_x^n| \, dx\, dt + \left|\int_0^T \int_0^L hv(u_x^n - u_x) \, dx\, dt\right| \\
&\le \left(\int_0^T \int_0^L |(h^n - h)v|^2 \, dx\, dt\right)^{\frac{1}{2}} \left(\int_0^T \int_0^L |u_x^n|^2 \, dx\, dt\right)^{\frac{1}{2}} \\
&\quad + \left|\int_0^T \int_0^L hv(u_x^n - u_x) \, dx\, dt\right| \\
&\le C_1 \left(\int_0^T \int_0^L |h^n - h|^2 \, dx\, dt\right)^{\frac{1}{2}} \left(\int_0^T \int_0^L |u_x^n|^2 \, dx\, dt\right)^{\frac{1}{2}} \\
&\quad + \left|\int_0^T \int_0^L hv(u_x^n - u_x) \, dx\, dt\right| \\
&\to 0 \quad \text{as } n \to \infty,
\end{aligned}$$

where the C_1 depends on the L^∞ bound on v. The convergence of the first term depends on the L^2 strong convergence of the $\{h^n\}$ sequence and the L^2 uniform bound on the $\{u_x^n\}$ sequence. Since hv is in $L^2(Q)$, the second term converges by weak L^2 convergence of the $\{u_x^n\}$ sequence. Passing to the limit in the PDE satisfied by u^n, we obtain $u^* = u(h^*)$.

Next, we verify that h^* is an optimal control. Using the lower semicontinuity of the objective functional with respect to these weak convergences, we obtain

$$\begin{aligned}
J(h^*) &\le \inf_n \alpha \int_0^T \int_0^L u^n(x,t) \, dx + \inf_n \frac{\beta}{2} \int_0^T \int_0^L \left[(h^n)^2 + (h_t^n)^2 + (h_x^n)^2\right] \, dx\, dt \\
&= \lim_{n\to\infty} J(h^n) = \inf_h J(h).
\end{aligned}$$

Hence h^* is an optimal control. $\square$

Remark: The convergence of the term $h^n u_x^n$ with $h^n = h^n(x,t)$ is precisely the reason for having a H^1 control set. If the controls were only functions of space, then the L^∞ control sets and techniques of [**1, 4**] would work.

3. Characterization of an Optimal Control

By differentiating the objective functional $J(h)$ with respect to h, we derive the optimality system. We will obtain the characterization of an optimal control through an elliptic differential equation. Thus the optimality system will consist of two parabolic equations coupled with an elliptic equation.

LEMMA 3.1. *The solution map*

$$h \in U \longrightarrow u = u(h) \in W$$

is differentiable in the following sense:

$$\frac{u(h+\varepsilon l) - u(h)}{\varepsilon} \rightharpoonup \phi \quad \textit{weakly in } W$$

as $\varepsilon \to 0$, *where* $h + \varepsilon l \in U, l \in L^\infty(Q)$. *Moreover the sensitivity function* $\phi = \phi(h;l)$ *satisfies*

$$\begin{aligned}
\phi_t(x,t) &= a\phi_{xx} - h\phi_x - l u_x && \textit{in } Q = (0,L)\times(0,T) \\
\phi(x,0) &= 0 && \textit{when } t = 0 \\
\phi(0,t) &= 0 && \textit{when } x = 0 \\
\phi_x(L,t) &= 0 && \textit{when } x = L,
\end{aligned} \tag{3.1}$$

where $u = u(h)$.

PROOF. In order to get *a priori* estimates on the difference quotients, we do the following change of variables

$$w = e^{-\lambda t} u, \quad w^\varepsilon = e^{-\lambda t} u^\varepsilon,$$

where $u = u(h)$ and $u^\varepsilon = u(h+\varepsilon l)$. By considering the PDE solved by $w^\varepsilon - w$, we obtain (multiplying by test function $(w^\varepsilon - w)$)

$$\begin{aligned}
&\frac{1}{2}\int_0^L (w^\varepsilon - w)^2(x,T)\,dx + a\int_0^T\int_0^L ((w^\varepsilon - w)_x)^2\,dx\,dt + \lambda\int_0^T\int_0^L (w^\varepsilon - w)^2\,dx\,dt \\
&= -\int_0^T\int_0^L h(w^\varepsilon - w)_x(w^\varepsilon - w)\,dx\,dt - \varepsilon\int_0^T\int_0^L l w_x^\varepsilon (w^\varepsilon - w)\,dx\,dt \\
&\le \frac{a}{2}\int_0^T\int_0^L ((w^\varepsilon - w)_x)^2\,dx\,dt + \frac{1}{2a}\int_0^T\int_0^L (h(w^\varepsilon - w))^2\,dx\,dt \\
&\quad + \frac{\varepsilon^2}{4}\int_0^T\int_0^L (l w_x)^2\,dx\,dt + \int_0^T\int_0^L (w^\varepsilon - w)^2\,dx\,dt \\
&\le \frac{a}{2}\int_0^T\int_0^L ((w^\varepsilon - w)_x)^2\,dx\,dt + \frac{M^2}{2a}\int_0^T\int_0^L (w^\varepsilon - w)^2\,dx\,dt \\
&\quad + \frac{\varepsilon^2}{4}\int_0^T\int_0^L (l w_x)^2\,dx\,dt + \int_0^T\int_0^L (w^\varepsilon - w)^2\,dx\,dt + K_1\varepsilon^2, \qquad (3.2)
\end{aligned}$$

denoting the L^∞ bound on h by M. Using the facts that $l \in L^\infty, h \in H^1(Q) \subset C^{0,1/2}(\bar{Q})$, and $\|w_x^\varepsilon\|_{L^2}$ is uniformly bounded from (2.3), we obtain

$$\frac{a}{2}\int_0^T\int_0^L ((w^\varepsilon - w)_x)^2\,dx\,dt + \left(\lambda - \frac{M^2}{2a} - 1\right)\int_0^T\int_0^L (w^\varepsilon - w)^2\,dx\,dt \le K_1\varepsilon^2,$$

where K_1 depends on $\|l\|_\infty$ and $\|w_x^\varepsilon\|_{L^2}$. For large λ, we have

$$\left\|\frac{w^\varepsilon - w}{\varepsilon}\right\|_W \le K_2$$

or

$$\left\|\frac{u^\varepsilon - u}{\varepsilon}\right\|_W \le K_3.$$

Thus weak convergence to ϕ is achieved. We also have $\left(\dfrac{u^\varepsilon - u}{\varepsilon}\right)_t \rightharpoonup \phi_t$ weakly in $L^2(0,T;V^*)$ and $\dfrac{u^\varepsilon - u}{\varepsilon} \longrightarrow \phi$ strongly in $L^2(Q)$ as in Theorem 2.1. Hence, we conclude that ϕ satisfies (3.1). $\square$

Now we derive a PDE characterizing an optimal control by differentiating $J(h)$ with respect to h at an optimal control.

THEOREM 3.1. *Given an optimal control h in U, there exists a solution p in W to the adjoint problem*

$$\begin{aligned}
&-p_t = ap_{xx} + (hp)_x + 1 && \text{in } Q \\
&p(x,T) = 0 && \text{on } (0,L), t = T \\
&p(0,t) = 0 && \text{on } (0,T), x = 0 \\
&ap_x(L,t) + hp(L,t) = 0 && \text{on } (0,T), x = L.
\end{aligned} \tag{3.3}$$

Furthermore $h(x,t)$ satisfies the following PDE.

$$\begin{aligned}
&h_{xx} + h_{tt} = h - \tfrac{\alpha}{\beta} p u_x && \text{in } Q \\
&\tfrac{\partial h}{\partial \eta} = 0 && \partial Q,
\end{aligned} \tag{3.4}$$

where $u = u(h)$.

Remark: The weak form of (3.4) is derived by:

$$\begin{aligned}
&\int_0^T \int_0^L \left(-h_{xx} - h_{tt} + h - \frac{\alpha}{\beta} p u_x \right) l(x,t)\, dx\, dt \\
= &\int_0^T \int_0^L \left(h_t l_t + h_x l_x + hl - \frac{\alpha}{\beta} p u_x l \right)(x,t)\, dx\, dt \\
&- \int_0^L \left(l h_t(x,T) - l h_t(x,0) \right) dx - \int_0^T \left(l h_x(L,t) - l h_x(0,t) \right) dt \\
= &\int_0^T \int_0^L \left(h_t l_t + h_x l_x + hl - \frac{\alpha}{\beta} p u_x l \right)(x,t)\, dx\, dt \quad \text{(using the BC from (3.4))}.
\end{aligned}$$

PROOF. Suppose $h(x,t)$ is an optimal control. Let $l \in H^1(Q)$ such that $h + \varepsilon l \in U$ for $\varepsilon > 0$. The derivative of $J(h)$ with respect to h in the direction l satisfies

$$\begin{aligned}
0 &= \lim_{\varepsilon \to 0} \frac{J(h + \varepsilon l) - J(h)}{\varepsilon} \\
&= \lim_{\varepsilon \to 0} \Bigg[\alpha \int_0^T \int_0^L \frac{u^\varepsilon - u}{\varepsilon}(x,t)\, dx\, dt \\
&\quad + \frac{\beta}{2\varepsilon} \int_0^T \int_0^L \left\{ (h + \varepsilon l)^2 - h^2 + (h_t + \varepsilon l_t)^2 - h_t^2 + (h_x + \varepsilon l_x)^2 - h_x^2 \right\}(x,t)\, dx\, dt \Bigg] \\
&= \alpha \int_0^T \int_0^L \phi(x,t)\, dx\, dt + \beta \int_0^T \int_0^L \left[hl + h_t l_t + h_x l_x \right](x,t)\, dx\, dt.
\end{aligned} \tag{3.5}$$

Since there are no pointwise inequality constraints on our control set, we have this directional derivative being zero.

We used Lemma 3.1 for the weak convergence of $(u^\varepsilon - u)/\varepsilon$.

Let p be the solution of the adjoint problem (3.3) [**5**]; then substituting the weak form of (3.3) in (3.5) gives

$$\begin{aligned} 0 &= \alpha \int_0^T \int_0^L \left(-p_t\phi + a p_x \phi_x - (hp)_x \phi\right)(x,t)\,dx\,dt + \alpha \int_0^T hp\phi(L,t)\,dt \\ &\quad +\beta \int_0^T \int_0^L \left[hl + h_t l_t + h_x l_x\right](x,t)\,dx\,dt. \end{aligned}$$

Integrating by parts yields

$$\begin{aligned} 0 &= \alpha \left\{ \int_0^L -[\phi p]_0^T\,dx + a\int_0^T p_x\phi_x\,dx\,dt - \int_0^T [\phi hp]_0^L\,dt + \int_0^T hp\phi(L,t)\,dt \right. \\ &\quad \left. + \int_0^T \int_0^L \left(p\phi_t + hp\phi_x\right)\,dx\,dt \right\} + \beta \int_0^T \int_0^L \left[hl + h_t l_t + h_x l_x\right](x,t)\,dx\,dt \\ &= \alpha \left\{ \int_0^L \left[-\phi p(x,T) + \phi p(x,0)\right]dx - \int_0^T \left[(hp - hp)\phi(L,t) - hp\phi(0,t)\right]dt \right\} \\ &\quad +\alpha \int_0^T \int_0^L \left(p\phi_t + a\phi_x p_x + hp\phi_x\right)\,dx\,dt + \beta \int_0^T \int_0^L \left[hl + h_t l_t + h_x l_x\right](x,t)\,dx\,dt \\ &= \int_0^T \int_0^L \left(-\alpha p l u_x + \beta\left[hl + h_t l_t + h_x l_x\right]\right)(x,t)\,dx\,dt. \quad \text{(using (3.1) and (3.3))} \end{aligned}$$

We substituted the inhomogenous term $-lu_x$ from the sensitivity equation (3.1) to obtain the last equality. The last equality is the weak form of the PDE (3.4), that an optimal control h must satisfy. $\square$

Remark: Our optimality system consists of two parabolic PDEs and an elliptic PDE, (1.1), (3.3), and (3.4). The uniqueness of the optimal control is not known at this time.

Bibliography

[1] Chawla, S. and Lenhart, S. *Application of Optimal Control Theory to Bioremediation*, J. of Computational and Applied Mathematics, 114 (2000), 81-102.

[2] Evans, L. C., *Partial Differential Equations*, American Mathematical Society, Providence, 1998.

[3] Handagama, N. and Lenhart, S., *Optimal Control of a PDE/ODE System Modeling A Gas-Phase Bioreactor*, Mathematical Models in Medicinal and Health Sciences, 1997, 197-212.

[4] Lenhart, S., *Optimal Control of a Convective-Diffusive Fluid Problem*, Math. Models and Method in Appl. Sc., Vol 5, No. 2 (1995), 225-237.

[5] Lions, J. L. and Magenes, E., *Non-Homogeneous Boundary Value Problems and Applications I & II*, Springer-Verlag, New York, 1972.

[6] Yong, J., Li, X., *Optimal Control Theory for Infinite Dimensional Systems*, Birkhäuser, 1995.

Vita

Hem Raj Joshi was born in Shera, Baitadi, Nepal on May 18, 1963. He obtained his first masters degree in mathematics from Garhwal University, India in 1989 and second masters degree in Industrial Mathematics from University of Kaiserslautern, Germany in 1995. Hem Raj Joshi married Dr. Rashika Bhattarai in 1996. He was blessed with a son Umang Raj Joshi in 1999 and a daughter Imani Joshi in 2001. He taught mathematics at Tribhuwan University for five years and was tenured in 1997. In August 1997, he entered The University of Tennessee, Knoxville to pursue the Doctorate of Philosophy in Mathematics. The doctoral degree was received in May 2002 under Professor Suzanne Lenhart, whom he met at the University of Kaiserslautern, where she was visiting Professor for a semester in 1994. He accepted a post-doctoral position in the Ecology and Evolutionary Biology Department at the University of Tennessee, Knoxville.